

Lecture notes on NIP theories

Pierre Simon

August 21, 2012

Contents

1	Introduction	5
1.1	Preliminaries	9
1.1.1	Indiscernible sequences	10
2	Definitions and basic properties	13
2.1	NIP formulas	13
2.1.1	The strict order property	16
2.1.2	Counting types	17
2.1.3	NIP theories	18
2.2	Invariant types	20
2.2.1	Products and Morley sequences	22
2.2.2	Eventual types	23
2.2.3	Generically stable types	25
3	Honest definitions and applications	29
3.1	Honest definitions	30
3.2	Naming a submodel	32
3.3	Shrinking of indiscernible sequences	34
4	Strong dependence and dp-ranks	39
4.1	Mutually indiscernible sequences	39
4.2	Dp-ranks	42
4.3	Strongly dependent theories	48
5	Forking	53
5.1	Bounded equivalence relations	53
5.2	Forking	56
5.3	$\text{bdd}(A)$ -invariance	59

5.4	NTP_2 and the broom lemma	61
5.5	Strict non-forking	64
6	Finite combinatorics	69
6.1	VC-dimension	69
6.2	The (p, q) -theorem	74
6.3	Uniformity of honest definitions	76
7	Measures	79
7.1	Definitions and basic properties	79
7.2	Boundedness properties	83
7.3	Smooth measures	84
7.4	Invariant measures	86
7.5	Generically stable measures	90
A	Examples of NIP structures	97
A.1	Linear orders and trees	97
A.2	Valued fields	101
B	Stability	109
C	Probability theory	111

Chapter 1

Introduction

This text is an introduction to the study of NIP (or dependent) theories. It is meant to serve two purposes. The first is to present various aspects of NIP theories and give the reader the background material needed to understand almost any paper on the subject. The second is to advertise the use of honest definitions, in particular in establishing basic results, such as the so-called *shrinking of indiscernibles*. Thus although we claim no originality for the theorems presented here, a few proofs are new, mainly in chapters 3 and 4.

We have tried to give a *horizontal* exposition, covering different, sometimes unrelated, topics at the expense of exhaustivity. Thus no particular subject is dealt with in depth and only *low-level* results are included. The choices made reflect our own interests and are certainly very subjective. In particular, we say very little about algebraic structures and concentrate on combinatorial aspects. Overall, the style is concise, but hopefully all details of the proofs are given. A small number of facts are left to the reader as exercises, but only once or twice are they used later in the text.

The material included is based on the work of a number of model theorists. Credits are usually not given alongside each theorem, but are recorded at the end of the chapter.

We have included almost no preliminaries about model theory, thus we assume some familiarity with basic notions, in particular concerning compactness, types and indiscernibles sequences. Those prerequisites are exposed in various books such as that of Poizat [72], Marker [61] or Hodges [40]. Familiarity with stability theory is not at all required. We will however make a few remarks concerning stable formulas or structures inside the text. The basic facts that we need are summarized in Appendix B.

History of the subject

In his early works on classification theory, Shelah structured the landscape of first order theories by drawing dividing lines defined by the presence or absence of different combinatorial configurations. The most important one is that of stability. In fact, for some twenty years, pure model theory did not venture much outside of stable theories. Shelah discovered the independence property, when studying the possible behaviors for the function relating the size of a subset to the number of types over it. The class of theories lacking the independence property, or NIP theories, was studied very little in the beginnings, especially by Shelah and Poizat (see [72, Chapter 12] for an account of those early works).

As the years passed, various structures were identified as being NIP: most notably henselian valued fields of characteristic 0 with NIP residue field and ordered group (Delon [26]), the field \mathbb{Q}_p of p-adics ([14]) and ordered abelian groups ([36]). However, NIP theories were not studied *per se*. In [67], Pillay and Steinhorn, building on work of van den Dries, defined o-minimal theories as a framework for tame geometry. This has been a very active area of research ever since. Although it was noticed from the start that o-minimal theories lacked the independence property, very little use of this fact was made until recently. Nevertheless, o-minimal theories provide a wealth of interesting examples of NIP structures.

In the years 2000, the interest in NIP theories was rekindled and the subject has been expanding ever since. First Shelah initiated a systematic study which lead to a series of paper: [82], [83], [77], [85], [84]. Amongst other things, he established the basic properties of forking, generalized a theorem of Baisalov and Poizat on externally definable sets, defined some subclasses, so called “strongly-dependent” and “strongly⁺-dependent”. This work culminates in [84] with the proof that NIP theories have few types up to automorphism (over saturated models). Parallel to this work, Hrushovski, Peterzil and Pillay developed the theory of *measures* (a notion introduced by Keisler in [55]) in order to solve Pillay’s conjecture on definably compact groups in o-minimal theories.

A third line of research starts with the work of Hrushovski, Haskell and Macpherson on algebraically closed valued fields (ACVF) and in particular on *metastability* ([38]). This lead to Hrushovski and Loeser giving a model theoretic construction of Berkovich spaces in rigid geometry as spaces of stably-dominated types, which made an explicit use of the NIP property along with

the work on metastability.

Motivated by those results, a number of model theorists got interested in the subject and investigated NIP theories in various directions. We will present some in the course of this text and mention others at the end of this introduction.

Let us end this general introduction by mentioning where NIP sits with respect to other classes of theories. First, all stable theories are NIP, as are o-minimal and C -minimal theories. Another well-studied extension of stability is that of simple theories (see [95]), however it is in a sense orthogonal to NIP: a theory is both simple and NIP if and only if it is stable. Simple and NIP theories both belong to the wider class of NTP_2 theories (defined in Chapter 5).

Organization of this text

Aside from the introduction and appendices, the text is divided into 6 chapters, each one focussing on a specific topic. In Chapter 2, we present the classical theory as it was established by Shelah and Poizat. We first work formula-by-formula giving some equivalent definitions of NIP. The original counting types characterization is included, although it will not be used later in the text. Starting at Section 2.1.3 and throughout most of the text, we assume that our ambient theory T is NIP. That assumption will be dropped only for the first three sections of Chapter 5. Of course, most of the results could be established for an NIP *formula* (or type) inside a (possibly) independent theory, but for the sake of clarity we will not work at this level of generality. The end of Chapter 2 is concerned with invariant and generically stable types.

In Chapter 3 we define honest definitions. They serve as a substitute to definability of types in NIP theories. We use them to prove Shelah's theorem on expanding a model by externally definable sets and the very important results about shrinking of indiscernibles.

Chapter 4 deals with dp-rank and strong dependence. In the literature, one can find up to three different definitions of dp-rank, based on how one handles the problem of *almost finite*, non-finite rank. None of them is perfect, but we have decided to use the same convention as in Adler's paper [2] on burden, although we refrain from duplicating limit cardinals into κ_- and κ . Instead, we define when $\text{dp-rk}(p) < \kappa$, and it may happen that the dp-rank of a type is not defined (for example it can be $< \aleph_0$ but greater than all integers).

In Chapter 5, we study forking and dividing. The main results are $\text{bdd}(A)$ -

invariance of non-forking extension from Hrushovski and Pillay's [47], and Chernikov and Kaplan's theorem about equality of forking and dividing over models. The right context for this latter result is NTP_2 theories, but here again we assume NIP which slightly simplifies some proofs.

The last two chapters have a different flavor. In Chapter 6, we change the framework to that of finite combinatorics. We are concerned with families of finite VC-dimension over finite sets. The finite and infinite approaches come together to prove uniformity of honest definitions. In Chapter 7, the two frameworks are combined with the introduction of Keisler measures. Here we stop short of giving applications to fsg groups and compact domination which can be found in [47] and [42].

Finally, three appendices are included. The first one gives some algebraic examples and in particular records some facts about valued fields. Most of the proofs are omitted, but we explain how to show that those structures are NIP. The other two appendices are very short and collect results about stability theory and probability theory for reference in the text.

Further topics

Many topics are entirely omitted from those notes. We mention some here with pointers to papers that the interested reader may turn to. We restrict ourselves to topics in pure model theory, and of course, we are by no means exhaustive.

- Compact domination: We will define generically stable measure in Chapter 7. Further information concerning them can be found in [42] and [43]. It is proved in [47] that a definable compact group in an o-minimal theory is dominated in some sense by a compact group. The link with generically stable measures is made in [42] where another compact domination result concerning type spaces is stated.
- Shelah's type decomposition theorems: Shelah showed in [85] that one can decompose types over saturated models in a finitely satisfiable component and a "directed" part. He then refined this greatly in [84] to show that over saturated models, there are few types up to automorphism.
- Directionality: In [53], Kaplan and Shelah group NIP theories into three classes called small, medium and large directionality depending on the number of coheirs that a type can have.
- Extraction of indiscernibles: Shelah proves in [77] that in a strongly-

dependent theory, one can extract an infinite indiscernible sequence from any long enough sequence. In [84], similar statements are proved for theories of small and medium directionality. Counter-examples to the general case are given in [53] and [52].

- Computing VC-density: The problem of computing precisely the VC-density of formulas in various NIP theories is investigated in [8] and [9].

- Groups interpretable in NIP structures: Few general statements are known about groups interpretable in NIP structures. Kaplan and Shelah [51] study various chain conditions. In another direction, pseudofinite NIP groups are studied in [60] and omega-categorical groups in [56].

It is proved in [50] that NIP fields have no Artin-Schreier extensions.

- Notions of minimality: We will define in 4.25 the class of dp-minimal theories. It provides us with a general combinatorial framework to study o-minimal and C -minimal theories, amongst others. Dp-minimal theories are studied for example in [64], [28] and [88]. A more restrictive class is that of VC-minimal theories which is defined by Adler in [3] and investigated in [34] and [25].

- Dependent pairs: A *pair* is a structure M with an additional predicate $\mathbf{P}(x)$ naming a given subset. Various pairs of NIP structures have been studied: [92], [15], [16], [35], [17]. The paper [23] contains a general criterion to ensure that the pair is still NIP.

- Distal theory: A distal theory is an NIP theory which is in some sense “purely unstable”. This notion is defined and studied in [86]. The paper [22] also contains some facts about it.

Acknowledgments The material presented here was exposed during a series of lectures in Paris. I would like to thank all the participants of this seminar: Élisabeth Bouscaren, Zoé Chatzidakis, Pablo Cubides Kovacsics, Françoise Delon, Martin Hils, Samaria Montenegro, Françoise Point, Silvain Rideau, Patrick Simonetta. Special thanks to Martin Hils for a number of helpful comments on this text.

1.1 Preliminaries

We work with a complete, one-sorted, theory T in a language L . We have a monster model \mathcal{U} which is $\bar{\kappa}$ -saturated and homogeneous. A subset $A \subset \mathcal{U}$ is **small** if it is of size less than $\bar{\kappa}$.

We do not distinguish between points and tuples. If a is a tuple of size $|a|$, we will usually write $a \in A$ to mean $a \in A^{|a|}$. Similarly, x, y, z, \dots are used to denote tuples of variables.

If $A \subset \mathcal{U}$ is any set, and $\phi(x)$ is a formula, then $\phi(A) = \{a \in A^{|x|} : \mathcal{U} \models \phi(a)\}$. The set of types over A in the variable x is denoted by $S_x(A)$. We will often drop the x . If $p \in S_x(A)$, we might write p_x or $p(x)$ to emphasize that p is a type in the variable x . We say that a type p **concentrates** on a definable set $\phi(x)$ if $p \vdash \phi(x)$. If $A \subseteq B$ and p is a type over B , we denote by $p \upharpoonright A$, or by $p|_A$ the **restriction** of p to A .

We use the notation ϕ^0 to mean $\neg\phi$ and ϕ^1 to mean ϕ . If $\phi(x; y)$ is a formula, a **ϕ -type** over A is a maximal consistent set of formulas of the form $\phi(x; a)^\epsilon$, for $a \in A$ and $\epsilon \in \{0, 1\}$. The set of ϕ -types over A is denoted by $S_\phi(A)$.

A **global type**, is a type over \mathcal{U} .

1.1.1 Indiscernible sequences

We will typically denote sequences of tuples by $I = (a_i : i \in \mathcal{I})$ where \mathcal{I} is some linearly ordered set. The order on \mathcal{I} will be denoted by $<_{\mathcal{I}}$ or simply $<$ if no confusion arises. If $I = (a_i : i \in \mathcal{I})$ and $J = (b_j : j \in \mathcal{J})$, then we write the concatenation of I and J as $I + J$. It has I as initial segment and J as the complementary final segment. We use the notation (a) to denote the sequence which has a as unique element.

We say that the sequence I is **endless** if the indexing order \mathcal{I} has no last element.

Let Δ be a finite set of formulas and A a set of parameters. A (possibly finite) sequence $I = (a_i : i \in \mathcal{I})$ is Δ -indiscernible over A , if for every integer k and two increasing tuples $i_1 <_{\mathcal{I}} \dots <_{\mathcal{I}} i_k$ and $j_1 <_{\mathcal{I}} \dots <_{\mathcal{I}} j_k$, $b \in A$ and formula $\phi(x_1, \dots, x_k; y) \in \Delta$, we have $\phi(a_{i_1}, \dots, a_{i_k}; b) \leftrightarrow \phi(a_{j_1}, \dots, a_{j_k}; b)$.

An **indiscernible sequence** is an *infinite* sequence which is Δ -indiscernible for all Δ . If the sequence I is indiscernible over A , then we define the **EM-type** of I over A to be the sequence $(p^n : 1 \leq n < \omega)$ of types where $p^n = \text{tp}(a_{i_1}, \dots, a_{i_n}/A)$ for some (any) $i_1 <_{\mathcal{I}} \dots <_{\mathcal{I}} i_n$. If $A = \emptyset$, then we can omit it. We will write $I \equiv_A^{EM} J$ to mean that I and J are two A -indiscernible sequences having the same EM-type over A . It follows from compactness that if I is an indiscernible sequence and \mathcal{J} is any small infinite linear order, then we can find an indiscernible sequence J indexed by \mathcal{J} and of same EM-type as I .

A sequence I is **totally indiscernible** (or set indiscernible) if every permutation of it is indiscernible.

If $(b_j : j < \alpha)$ is any sequence and $I = (a_i : i \in \mathcal{I})$ is an indiscernible sequence, we say that I is **based** on the sequence $(b_i)_{i < \alpha}$ if for every $i_1 <_{\mathcal{I}} \dots <_{\mathcal{I}} i_n$ and formula $\phi(x_1, \dots, x_n) \in L$, there are $j_1 < \dots < j_n < \alpha$ such that $\phi(a_{i_1}, \dots, a_{i_n}) \leftrightarrow \phi(b_{j_1}, \dots, b_{j_n})$ holds. We say **based on $(b_i)_{i < \alpha}$ over A** if we allow $\phi(x_1, \dots, x_n)$ to range over $L(A)$.

Ramsey's theorem and compactness imply that if $(b_i)_{i < \alpha}$ is any sequence, then there is an indiscernible sequence based on it.

Chapter 2

Definitions and basic properties

In this chapter, we introduce the basic objects of our study. We first define the notion of an NIP formula. The combinatorial definition is not very handy, and we give an equivalent characterization involving indiscernible sequences which is the one we will most often use. We then define NIP theories as theories in which all formulas are NIP. From that point on, we restrict our attention to that class of theories. We study invariant types and Morley sequences. Finally, we define the notion of a *generically stable* type and give some properties.

To illustrate the notions considered, we prove some results on definable groups in NIP theories: the Baldwin-Saxl theorem, and Shelah's theorem on existence of infinite definable abelian subgroups.

2.1 NIP formulas

Definition 2.1. A partitioned formula $\phi(x; y)$ is NIP if we cannot find an infinite set A of $|x|$ -tuples such that:

(IP) $_{\phi, A}$ for all $A_0 \subseteq A$, there is some b_{A_0} of size $|y|$ such that

$$\phi(A; b_{A_0}) = A_0.$$

If a formula is not NIP, we say that it has IP.

Remark 2.2. If $\phi(x; y)$ is NIP, then by compactness, there is some N such that (IP) $_{\phi, A}$ does not hold for any A of size at least N .

The maximal integer N for which there is some A of size N such that (IP) $_{\phi, A}$ holds is called the *VC-dimension* of ϕ . If there is no such integer (*i.e.*, the formula ϕ has IP), then we say that the VC-dimension is infinite.

EXAMPLE 2.3.

• Let $T = DLO$ be the theory of dense linear orders with no endpoints. Then the formula $\phi(x; y) = x \leq y$ is NIP of VC-dimension 1. Indeed, if we have $a_1 < a_2$, then we cannot find some $b_{\{a_2\}}$ such that

$$\mathcal{U} \models \neg\phi(a_1; b_{\{a_2\}}) \wedge \phi(a_2; b_{\{a_2\}}).$$

- If $\phi(x; y)$ is a stable formula, then it is NIP.
- If T is the theory of arithmetic, then the formula $\phi(x; y) = “x \text{ divides } y”$ has IP. To see this, take any $N \in \mathbb{N}$ and $A = \{p_0, \dots, p_{N-1}\}$ a set of distinct prime numbers. Then for any $A_0 \subseteq A$, set b_{A_0} to be $\prod_{a \in A_0} a$. We have $\models \phi(a, b_{A_0}) \iff a \in A_0$. This shows that $\phi(x; y)$ has infinite VC-dimension.
- If T is the random graph in the language $L = \{R\}$, then the formula $\phi(x; y) = xRy$ has IP.
- If T is a theory of an infinite boolean algebra, in the language $\{\leq, \cup, \cap\}$, then the formula $x \leq y$ has IP.

Lemma 2.4. *The formula $\phi(x; y)$ is NIP if and only if $\phi^{opp}(x; y) = \phi(y; x)$ is NIP.*

Proof. Assume that $\phi(x; y)$ has IP. Then by compactness, we can find some $A = \{a_i : i \in \mathfrak{P}(\omega)\}$ such that $(IP)_{\phi, A}$ holds, witnessed by tuples b_{A_0} , $A_0 \subseteq A$. Let $B = \{b_j : j \in \omega\}$ where $b_j := b_{A_j}$ and $A_j := \{D \subseteq \omega : j \in D\}$. Then for any $B_0 \subseteq B$, we have

$$\models \phi(a_{B_0}, b_j) \iff j \in B_0.$$

This shows that $(IP)_{\phi^{opp}, B}$ holds. Therefore ϕ^{opp} has IP. Note however that ϕ^{opp} need not have the same VC-dimension as ϕ . \square

We now give an equivalent characterization of NIP which is the one we will most often use.

Lemma 2.5. *The formula $\phi(x; y)$ has IP if and only if there is an indiscernible sequence $(a_i : i < \omega)$ and a tuple b such that*

$$\models \phi(a_i; b) \iff i \text{ is even}.$$

Proof. (\Leftarrow): Assume that there is a sequence $(a_i : i < \omega)$ and a tuple b as above. Let $A \subseteq \omega$. We show that there is some b_A such that $\phi(a_i; b_A)$ holds if and only if $i \in A$. We can find an increasing one-to-one map $\tau : \omega \rightarrow \omega$ such that for all $i \in \omega$, $\tau(i)$ is even if and only if i is in A . Then the map

sending b_i to $b_{\tau(i)}$, for all $i < \omega$ is a partial isomorphism. It extends to a global automorphism σ . Then take $a_A = \sigma^{-1}(a)$.

(\Rightarrow): Assume that $\phi(x; y)$ has IP. Let A be an infinite set of $|x|$ -tuples such that $(\text{IP})_{\phi, A}$ holds. By compactness we may assume that A is very large and then by Erdős-Rado, we can find some indiscernible sequence $I = (c_i : i < \omega)$ of $|x|$ -tuples such that for any n -tuple $\bar{c} = (c_{i_1}, \dots, c_{i_n})$ from I there is $\bar{a} = (a_1, \dots, a_n)$ from A having the same type as \bar{c} . This implies that $(\text{IP})_{\phi, I}$ holds. In particular, there is b such that $\phi(c_i; b)$ holds if and only if i is even. \square

It follows by compactness that if $\phi(x; y)$ is NIP, then there is a finite set Δ of formulas and an integer N such that the following does not exist:

- $(a_i : i < N)$ a Δ -indiscernible sequence of $|x|$ -tuples;
- b a $|y|$ -tuple, such that $\neg(\phi(a_i; b) \leftrightarrow \phi(a_{i+1}; b))$ holds for $i < N - 1$.

Let $I = (a_i : i \in \mathcal{I})$ be an indiscernible sequence, and let $\phi(x; b) \in L(\mathcal{U})$ be a formula. Then there is a maximal integer n such that we can find $i_0 <_{\mathcal{I}} \dots <_{\mathcal{I}} i_n$ with $\neg(\phi(a_{i_k}; b) \leftrightarrow \phi(a_{i_{k+1}}; b))$ for all $k < n$. We call such an n the number of alternations of $\phi(x; b)$ on the sequence I and write it as $\text{alt}(\phi(x; b), I)$. We let $\text{alt}(\phi(x; y))$ denote the maximum value of $\text{alt}(\phi(x; b), I)$, for b ranging in \mathcal{U} and I ranging over all indiscernible sequences. Note that this maximum exists and is bounded by the number N of the previous paragraph.

Proposition 2.6. *The formula $\phi(x; y)$ is NIP if and only if for any indiscernible sequence $(a_i : i \in \mathcal{I})$ and tuple b , there is some end segment $\mathcal{I}_0 \subseteq \mathcal{I}$ and $\eta \in \{0, 1\}$ such that $\phi(a_i; b)^\eta$ holds for any $i \in \mathcal{I}_0$.*

Proof. If \mathcal{I} has a last elements i_0 , let $\mathcal{I}_0 = \{i_0\}$. Otherwise, this follows immediately from Lemma 2.4. \square

Lemma 2.7. *A boolean combination of NIP formulas is NIP.*

Proof. It is clear from the definition that the negation of an NIP formula is NIP.

Let $\phi(x; y)$ and $\psi(x; y)$ be two NIP formulas and we want to show that $\theta(x; y) \equiv \phi(x; y) \wedge \psi(x; y)$ is NIP. We use the criterion from Proposition 2.6. Let $(a_i : i \in \mathcal{I})$ be an indiscernible sequence of $|x|$ -tuples and let b be a $|y|$ -tuple. Let $\mathcal{I}_\phi \subset \mathcal{I}$ be an end segment such that $\phi(a_i; b) \leftrightarrow \phi(a_j; b)$ holds for $i, j \in \mathcal{I}_\phi$. Define \mathcal{I}_ψ similarly. Then let $\mathcal{I}_0 = \mathcal{I}_\phi \cap \mathcal{I}_\psi$. Then \mathcal{I}_0 is an end segment and we have $\theta(a_i; b) \leftrightarrow \theta(a_j; b)$ for $i, j \in \mathcal{I}_0$. This shows that $\theta(x; y)$ is NIP. \square

2.1.1 The strict order property

Definition 2.8. We say that a formula $\phi(x; y)$ has the strict order property (SOP) if there exists a sequence $(b_i : i < \omega)$ of $|y|$ -tuples such that for all $i < \omega$,

$$\phi(\mathcal{U}; b_i) \subsetneq \phi(\mathcal{U}; b_{i+1}).$$

Remark 2.9. For the theory T , it is equivalent to say that some formula has SOP, or to say that there is, in T^{eq} , a partial definable order with infinite chains. (If $\phi(x; y)$ has SOP, then the formula $\psi(y_1, y_2) = (\forall x)\phi(x; y_1) \rightarrow \phi(x; y_2)$ defines an order with infinite chains.)

Theorem 2.10. Assume that T is unstable. Then at least one of the following holds:

- there is a formula $\phi(x; y)$ which has IP;
- there is a formula $\phi(x; y)$ which has SOP.

Proof. Let $\phi(x; y)$ be unstable and NIP. There is some indiscernible sequence $I = (a_i : i < \omega)$ such that for every $N < \omega$, we can find b_N such that $\phi(a_i; b)$ holds if and only if $i < N$. By NIP, there is some integer n and $\eta : n \rightarrow \{0, 1\}$ such that

$$\bigwedge_{i < n} \phi(a_i; y)^{\eta(i)}$$

is inconsistent. Starting with that formula, we change one by one instances of $\neg\phi(a_i; y) \wedge \phi(a_{i+1}; y)$ to $\phi(a_i; y) \wedge \neg\phi(a_{i+1}; y)$. In the end, we arrive at a formula of the form $\bigwedge_{i < N} \phi(a_i; y) \wedge \bigwedge_{N \leq i < n} \phi(a_i; y)$. The tuple b_N satisfies that formula. There is therefore one step in the process in which we change an inconsistent formula to a consistent one. Namely, there is some $i_0 < n$, $\eta_0 : n \rightarrow \{0, 1\}$ such that

$$\bigwedge_{i < i_0} \phi(a_i; y)^{\eta(i)} \wedge \neg\phi(a_{i_0}; y) \wedge \phi(a_{i_0+1}; y) \wedge \bigwedge_{i_0+1 < i < n} \phi(a_i; y)^{\eta(i)}$$

is inconsistent, but

$$\bigwedge_{i < i_0} \phi(a_i; y)^{\eta(i)} \wedge \phi(a_{i_0}; y) \wedge \neg\phi(a_{i_0+1}; y) \wedge \bigwedge_{i_0+1 < i < n} \phi(a_i; y)^{\eta(i)}$$

is consistent. Write those formulas as $\theta(\bar{a}; y) \wedge \neg\phi(a_{i_0}; y) \wedge \phi(a_{i_0+1}; y)$ and $\theta(\bar{a}; y) \wedge \phi(a_{i_0}; y) \wedge \neg\phi(a_{i_0+1}; y)$ respectively.

Increase the sequence I to $I' = (a_i : i \in \mathbb{Q})$, indexed by \mathbb{Q} . Then, for $i_0 \leq i, i' \leq i_0 + 1$, we have: $(\exists y)\theta(\bar{a}; y) \wedge \phi(a_i; y) \wedge \neg\phi(a_{i'}, y)$ holds if and only if $i < i'$. It follows that the formula $\psi(y; x) = \theta(\bar{a}; y) \wedge \phi(x; y)$ has the strict order property. \square

2.1.2 Counting types

We give another characterization of NIP by counting types. This characterization however relies on set-theoretic assumptions, and will not be used later in this text.

Fix some formula $\phi(x; y)$. The *stability function* for ϕ is the function g_ϕ defined on cardinals by $g_\phi(\kappa) = \sup\{|S_\phi(A)| : A \text{ of size } \kappa\}$. Recall that $S_\phi(A)$ denotes the set of ϕ -types over A .

If $\phi(x; y)$ is stable, then there is some integer n such that $g_\phi(\kappa) \leq \kappa^n$ for every (finite or infinite) cardinal κ . Conversely, if $g_\phi(\kappa) = \kappa$ for some infinite cardinal κ , then ϕ is stable. (See Appendix B.)

If $\phi(x; y)$ has IP, then easily, we have $g_\phi(\kappa) = 2^\kappa$ for every cardinal κ . Conversely, it follows from the definition that if $g_\phi(\kappa) = 2^\kappa$ for every *finite* κ , then ϕ has IP.

The case of an unstable NIP formula $\phi(x; y)$ is trickier. We will see in Chapter 6, that there is n such that for any *finite* cardinal κ , we have $g_\phi(\kappa) \leq \kappa^n$. If κ is infinite and $2^\kappa = \kappa^+$, then as ϕ is unstable, we must have $g_\phi(\kappa) = 2^\kappa$ so there is no hope of separating IP from NIP by counting types over infinite sets without extra set-theoretic assumptions.

Definition 2.11. For a cardinal λ , we define $\text{ded}(\lambda) = \sup\{\kappa : \text{there is a linear order of size } \kappa \text{ which has a dense subset of size } \lambda\}$.

Proposition 2.12. *Let $\phi(x; y)$ be a formula. Assume that there is some infinite set A with $|S_\phi(A)| > \text{ded}(|A|)$, then $\phi(x; y)$ has IP.*

Proof. Assume that $|S_\phi(A)| > \text{ded}(|A|)$ and that A is chosen such that $\mu = |A|$ is minimal. Let $\lambda = \text{ded}(|A|)^+$. Enumerate A as $\{a_i : i < \mu\}$ and for $i < \mu$, set $A_i = \{a_j : j < i\}$. For each $i < \mu$, we then have $|S_\phi(A_i)| \leq \text{ded}(|A_i|) < \lambda$.

Define the following sets:

$$\begin{aligned} S_i &= \{p \in S_\phi(A_i) : p \text{ has } \geq \lambda \text{ extensions to } S_\phi(A)\}, \text{ for } i < \mu; \\ S_\mu &= \{p \in S_\phi(A) : p \upharpoonright A_i \in S_i \text{ for all } i < \mu\}. \end{aligned}$$

Note that for every $i < \mu$, as $|S_\phi(A_i)| < \lambda$, we have $|S_\phi(A_i) \setminus S_i| < \lambda$ and as each type from that set has less than λ extensions to a type over A , the cardinality of $\{p \in S_\phi(A) : p \upharpoonright A_i \notin S_i\}$ is less than λ . Summing over $i < \mu$, we see that

$$|S_\phi(A) \setminus S_\mu| < \lambda.$$

It follows that every type in S_i has at least λ extensions to a type in S_μ .

Let $S_{<\mu} = \bigcup_{i<\mu} S_i$ and $S_{\leq\mu} = S_{<\mu} \cup S_\mu$. We define a linear order on $S_{\leq\mu}$ in the following way. For $p, q \in S_{\leq\mu}$, if $p \subseteq q$ (resp. $q \subseteq p$), we set $p \leq q$ (resp. $q \leq p$). Otherwise, let $i < \mu$ be maximal such that $p \upharpoonright S_i = q \upharpoonright S_i$. Then set $p < q$ if $p \vdash \neg\phi(x; a_i)$ (which implies $q \vdash \phi(x; a_i)$) and $p > q$ otherwise. We leave it to the reader to check that this indeed defines a linear order with $S_{<\mu}$ as a dense subset.

We now show:

(*) For all $n < \omega$ and $q \in S_{<\mu}$, there are tuples $b_q^0, \dots, b_q^{n-1} \in A$ such that for every $\eta : n \rightarrow \{0, 1\}$, the type $q(x) \wedge \bigwedge_{k < n} \phi(x; b_q^k)^{\eta(k)}$ is consistent.

This will imply that $\phi(x; y)$ has the independence property. We show (*) by induction on n . For $n = 0$, there is nothing to prove.

Assume the result is known for n , and we prove it for $n + 1$. Let $q \in S_i$ for some $i < \mu$. Define $S_q = \{p \in S_{\leq\mu} : p \upharpoonright S_i = q\}$. Using the order defined above, we see that $S_q \cap S_{<\mu}$ is a dense subset of S_q . We know that $|S_q| \geq \lambda$, therefore by definition of ded , $|S_q \cap S_{<\mu}| > \mu$. It follows that there is some $i < \mu$ such that $|S_q \cap S_i| > \mu$. The induction hypothesis gives, for every type $p \in S_q \cap S_i$, a tuple $(b_p^0, \dots, b_p^{n-1})$. We can find two distinct types $p_1, p_2 \in S_q \cap S_i$ for which the corresponding tuples are the same, equal to some (b_0, \dots, b_{n-1}) . Let $b_n \in A_i$ be such that $p_1 \vdash \phi(x; b_n)$ and $p_2 \vdash \neg\phi(x; b_n)$ (exchanging the roles of p_1 and p_2 if necessary). Then for every $\eta : n + 1 \rightarrow \{0, 1\}$, the partial type $q(x) \wedge \bigwedge_{k \leq n} \phi(x; b_k)^{\eta(k)}$ is consistent. This finishes the induction step, and the proof. \square

For any infinite cardinal κ , we have $\kappa < \text{ded}(\kappa) \leq 2^\kappa$. Hence if $2^\kappa = \kappa^+$, then $\text{ded}(\kappa) = 2^\kappa$ and the hypothesis of the previous proposition cannot be satisfied. However, Mitchell has shown in [63] that if $\text{cf}(\kappa) > \aleph_0$, then there is a cardinal-preserving forcing extension of the set-theoretical universe on which $\text{ded}(\kappa) < 2^\kappa$. In a such an extension, a formula $\phi(x; y)$ is IP if and only if $g_\phi(\kappa) = 2^\kappa$.

2.1.3 NIP theories

Definition 2.13. The theory T is NIP if all formulas $\phi(x; y)$ are NIP.

Proposition 2.14. Assume that all formulas $\phi(x; y)$, where $|x| = 1$ are NIP, then T is NIP.

Proof. Assume that all formulas $\phi(x; y)$ where $|x| = 1$ are NIP.

Claim : Let $(b_i : i < |T|^+)$ be an indiscernible sequence, and let $a \in \mathcal{U}$, $|a| = 1$. Then there is some $\alpha < |T|^+$ such that the sequence $(b_i : \alpha < i < |T|^+)$ is indiscernible over a .

Proof: Otherwise, there is some formula $\delta(x; \bar{y}_1, \dots, \bar{y}_k)$ such that for all $\alpha < |T|^+$, we can find some

$$\alpha < i_1 < \dots < i_k < |T|^+ \quad \text{and} \quad \alpha \leq j_1 < \dots < j_k < |T|^+$$

such that $\delta(a, \bar{b}_{i_1}, \dots, \bar{b}_{i_k}) \wedge \neg \delta(a, \bar{b}_{j_1}, \dots, \bar{b}_{j_k})$ holds. It follows that we can find a sequence $I = (i_1^l \wedge \dots \wedge i_k^l : l < \omega)$ such that $i_1^l < \dots < i_k^l < i_1^{l+1}$ for all $l < \omega$ and we have $\delta(a, \bar{b}_{i_1^l}, \dots, \bar{b}_{i_k^l})$ if and only if l is even. As the sequence I is indiscernible, this contradicts NIP.

Now let $\phi(\bar{x}; \bar{y})$ be any formula, with $\bar{x} = (x_1, \dots, x_n)$. Let $(\bar{b}_i : i < |T|^+)$ be an indiscernible sequence and let $\bar{a} = (a_1, \dots, a_n)$ be any tuple. By the claim, there is some $\alpha_1 < |T|^+$ such that the sequence $(\bar{b}_i : \alpha_1 \leq i < |T|^+)$ is indiscernible over a_1 . This implies that the sequence $(\bar{b}_i \hat{\ } a_1 : \alpha_1 < i < |T|^+)$ is indiscernible. Therefore another application of the claim gives some $\alpha_2 < |T|^+$ such that the sequence $(\bar{b}_i \hat{\ } a_1 : \alpha_2 < i < |T|^+)$ is indiscernible over a_2 . Iterating, we find $\alpha_n < |T|^+$ such that $(\bar{b}_i \hat{\ } a_1 \dots a_n : \alpha_n < i < |T|^+)$ is indiscernible. This implies that the truth value of $\phi(\bar{a}; \bar{b}_i)$ is constant for $i > \alpha_n$. Therefore the formula $\phi(x; y)$ is NIP. \square

EXAMPLE 2.15. The following theories are NIP:

- any stable theory;
- any o-minimal theory;
- any C -minimal theory, for example ACVF, the theory of algebraically closed valued fields;
- $Th(\mathbb{Q}_p)$;
- any pure tree.

See Appendix A for the definitions of some of those theories and extra details.

We give an example of how the definition of the independence property can be used.

Theorem 2.16 (Baldwin-Saxl). *Let G be a group definable in an NIP theory T . Let H_a be a uniformly definable family of subgroups of G . Then there is an integer N such that for any finite intersection $\bigcap_{a \in A} H_a$, there is a subset $A_0 \subset A$ of size N with $\bigcap_{a \in A} H_a = \bigcap_{a \in A_0} H_a$.*

Proof. The fact that H_a is a uniformly definable family of subgroups means that there is a formula $\phi(x; y)$ such that for any a , $\phi(x; a)$ defines a subgroup of G . Take any integer N , and assume that the conclusion of the theorem does not hold for N . Then we can find some set $A = \{a_0, \dots, a_N\}$ of parameters such that for every $k \leq N$, we have $\bigcap_{a \in A \setminus \{a_k\}} H_a \neq \bigcap_{a \in A} H_a$. Let $K_k = \bigcap_{a \in A \setminus \{a_k\}} H_a$ and $K = \bigcap_{a \in A} H_a$. For every $k \leq N$, pick a point $c_k \in K_k \setminus K$. For $B \subset N + 1$, define $c_B = \prod_{k \in B} c_k$, where the product is in the sense of the group G . Then we have

$$c_B \in H_k \iff k \notin B.$$

This shows that the formula $\psi(y; x) := \phi(x; y)$ has VC-dimension at least N . Therefore there is a maximal such N . \square

2.2 Invariant types

Assumption: From now on, we assume that the theory T is NIP (except in Proposition 2.31 where we give an additional characterization of NIP).

Definition 2.17 (Invariant type). Let $A \subset \mathcal{U}$ and $p \in S_x(\mathcal{U})$. We say that p is A -invariant if $\sigma p = p$ for any $\sigma \in \text{Aut}(\mathcal{U}/A)$.

We say that p is invariant, if it is A -invariant for some small $A \subset \mathcal{U}$.

Another way to phrase the definition is to say that p is A -invariant if for every formula $\phi(x; y)$ and tuples $b, b' \in \mathcal{U}$, if $b \equiv_A b'$, then

$$p \vdash \phi(x; b) \iff p \vdash \phi(x; b').$$

EXAMPLE 2.18. A global type $p \in S_x(\mathcal{U})$ is said to be *definable* over A if for every formula $\phi(x; y)$ without parameters, there is some formula $d\phi(y) \in L(A)$ such that $p \vdash \phi(x; b) \iff b \models d\phi(y)$ for all $b \in \mathcal{U}$. Any A -definable type is clearly A -invariant. We will say that a global type p is definable if it is A -definable for some A .

EXAMPLE 2.19. A type p is said to be *finitely satisfiable* in a set A if for every formula $\phi(x; b) \in p$, there is $a \in A$ such that $\phi(a; b)$ holds. If $p \in S_x(\mathcal{U})$ is finitely satisfiable in A , then it is A -invariant. A global type p is finitely satisfiable if it is finitely satisfiable in A for some small A . We present two constructions to obtain such types.

(1) Let $A \subset \mathcal{U}$ be any set and let \mathcal{D} be an ultrafilter on $A^{|x|}$. We define $p_{\mathcal{D}} \in S_x(\mathcal{U})$ by:

$$p_{\mathcal{D}} \vdash \phi(x; b) \iff \phi(A; b) \in \mathcal{D},$$

for every formula $\phi(x; b) \in L(\mathcal{U})$. Every type finitely satisfiable in A is equal to $p_{\mathcal{D}}$ for some ultrafilter \mathcal{D} . In particular note that if we take \mathcal{D} to be a principal ultrafilter, then we obtain a realized type.

(2) Let $I = (a_i : i \in \mathcal{I})$ be an indiscernible sequence. Then by Proposition 2.6, the sequence $(\text{tp}(a_i/\mathcal{U}) : i \in \mathcal{I})$ converges in $S(\mathcal{U})$ to some type, called the *limit type* of the sequence I , denoted by $\lim(I)$. This type is finitely satisfiable in I , and indeed in any cofinal subsequence of I .

If $p_0 \in S_x(M)$ is a type, then a *coheir* of p_0 is a global extension p which is finitely satisfiable in M . Such a coheir always exists as can be seen by extending $\{\phi(M; a) : \phi(x; a) \in p_0\}$ into an ultrafilter \mathcal{D} on $M^{|x|}$ and using considering $p_{\mathcal{D}}$.

Remark 2.20. Let $p \in S(\mathcal{U})$ be an A -invariant type. Let $M \supset A$ be $|A|^+$ -saturated and define $q = p|_M$. Then q has a unique extension to a type over \mathcal{U} which is A -invariant. (Let \tilde{p} be such an extension, and consider a formula $\phi(x; b) \in L(\mathcal{U})$, then there is $b' \in M$ such that $b \equiv_A b'$. By A -invariance, we have $\tilde{p} \vdash \phi(x; b) \iff \tilde{p} \vdash \phi(x; b') \iff p \vdash \phi(x; b) \iff p \vdash \phi(x; b')$.)

In particular, if $p \in S(\mathcal{U})$ is an invariant type, and if $V \supset \mathcal{U}$, then there is a unique extension of p to a type over V which is invariant over some small $A \subset \mathcal{U}$. We will denote that extension by $p|_V$.

Lemma 2.21. *Let $p \in S(\mathcal{U})$ be an A -invariant type, then:*

- if p is definable, it is A -definable;
- if p is finitely satisfiable, it is finitely satisfiable in any model $M \supset A$.

Proof. Assume that p is definable. Let $\phi(x; y)$ be any formula and let $d\phi(y) \in L(\mathcal{U})$ be such that for any $b \in \mathcal{U}$, $p \vdash \phi(x; b) \iff b \models d\phi(y)$. Then the definable set $d\phi(y)$ is invariant under all automorphisms that fix A pointwise. It follows that it is A -definable.

Assume now that p is finitely satisfiable in some small model N . Let M be any small model containing A . Let $\phi(x; d) \in L(\mathcal{U})$ be any formula and let N_1 realize a coheir of $\text{tp}(N/M)$ over Md . Then by invariance, p is finitely satisfiable in N_1 and in particular $\phi(N_1; d)$ is non-empty. By the coheir hypothesis, $\phi(M; d)$ is non-empty. Therefore p is finitely satisfiable in M . \square

2.2.1 Products and Morley sequences

Let $p_x, q_y \in S(\mathcal{U})$ be two invariant types. We define the type $p_x \otimes q_y \in S_{xy}(\mathcal{U})$ as $\text{tp}(a, b/\mathcal{U})$ where $b \models q_y$ and $a \models p|_{\mathcal{U}b}$.

Another way to say the same thing, if p and q are A -invariant: $p_x \otimes q_y \vdash \phi(x; y)$ if $p_x \vdash \phi(x; b)$ for some (any) $b \in \mathcal{U}$ with $b \models q_y|_A$.

The relation \otimes is associative, but in general not commutative.

More generally, if p_x is M -invariant, $N \supset M$ and $q_y \in S(N)$ is any type, then we can define $p_x \otimes q_y \in S_{xy}(N)$ as $\text{tp}(a, b/N)$ where $b \models q_y$ and $a \models p \upharpoonright Nb$.

EXAMPLE 2.22. Let T be DLO, and take $p = q$ to be the type at $+\infty$. Then $p_x \otimes q_y \vdash x > y$ whereas $q_y \otimes p_x \vdash x < y$.

Lemma 2.23. *If p_x is a global definable type and q_y is finitely satisfiable, then $p_x \otimes q_y = q_y \otimes p_x$.*

Proof. Let $M \prec \mathcal{U}$ be a model such that p is definable over M and q is finitely satisfiable in M . Let $\phi(x, y; d) \in L(\mathcal{U})$. By definability of p , there is a formula $d\phi(y; z) \in L(M)$ such that for every $b, d' \in \mathcal{U}$, we have $p \vdash \phi(x, b; d') \iff b \models d\phi(x; d')$. Let $\eta \in \{0, 1\}$ such that $q \vdash d\phi(x; d')^\eta$.

Without loss, assume that $q_y \otimes p_x \vdash \phi(x, y; d)$. Let $a \models p$. Then $q_y|_{\mathcal{U}a} \vdash \phi(a, y; d)$ and as q is finitely satisfiable, there is $b \in M$ such that $b \models \phi(a, y; d) \wedge d\phi(y; d)^\eta$. Then $p \vdash \phi(x, b; d) \wedge d\phi(b; d)^\eta$ and by definition of $d\phi$, this implies that $\eta = 1$.

Therefore $q \vdash d\phi(y; d)$ and it follows that $p_x \otimes q_y \vdash \phi(x, y; d)$. \square

If p_x is an A -invariant type, we define by induction for $n \in \mathbb{N}^*$:

$$p_{x_0}^{(1)} = p_{x_0} \quad \text{and} \quad p_{x_0, \dots, x_n}^{(n+1)} = p_{x_n} \otimes p_{x_0, \dots, x_{n-1}}^{(n)}.$$

Let also $p_{x_0, x_1, \dots}^{(\omega)} = \bigcup p^{(n)}$. For any $N \supset A$, a realization $(a_i : i < \omega)$ of $p^{(\omega)}|_N$ is called a *Morley sequence* of p over N (indexed by ω). Such a sequence $(a_i : i < \omega)$ is indiscernible over N (using associativity, we can check by induction on n that for any $i_1, \dots, i_n \in \omega$, we have $\text{tp}(a_{i_1}, \dots, a_{i_n}/N) = p^{(n)}|_N$).

More generally, any sequence indiscernible over N whose EM-type over N is given by $(p^{(n)}|_N : 1 \leq n < \omega)$ is called a Morley sequence of p over N .

EXERCISE 2.24. *If p is a global A -definable type (resp. finitely satisfiable in A), then $p^{(\omega)}$ is also A -definable (resp. finitely satisfiable in A).*

Let $I = (a_t : t \in \mathcal{I})$ be an endless indiscernible sequence. We have defined the limit type $p = \lim(I)$ as a global invariant type. Let \mathcal{J} be any linear order and $(c_t : t \in \mathcal{J})$ be a sequence such that $c_t \models p \upharpoonright I + c_{>t}$. So J is a Morley sequence of p over I indexed backwards. Then $I + J$ is indiscernible.

An application to definable groups

Lemma 2.25. *Let G be a definable group. Let p be a non-realized invariant type concentrating on G and $(a_i : i < \omega)$ a Morley sequence of p . Assume that for every $i, j < \omega$, $a_i a_j = a_j a_i$, then there is an infinite abelian definable subgroup H of G , such that p concentrates on H .*

Proof. First, we note that it is enough to find some infinite definable set $D \subset G$ such that p concentrates on D and any two elements of D commute. Indeed, if we have such a set D , we let $H = C(C(D))$, where $C(X) = \{y \in G : yx = xy \text{ for all } x \in X\}$. Then H is a definable abelian subgroup containing D .

We try to build by induction a sequence $(a_n b_n : n < \omega)$ such that $a_n \models p \upharpoonright a_{<n} b_{<n}$, $b_n \models p \upharpoonright a_{<n} b_{<n}$ and $a_n b_n \neq b_n a_n$. Assume that we succeed. Then by hypothesis, $a_n b_m = b_m a_n$ for $n \neq m$. For any $I \subset \omega$ finite, define $b_I = \prod_{n \in I} b_n$. We have thus, $a_n b_I = b_I a_n$ if and only if $n \notin I$. This contradicts NIP.

Therefore the construction must stop and for some n , we cannot construct a_n, b_n as required. Then if we let $q = p \upharpoonright a_{<n} b_{<n}$, we have $q(x) \wedge q(y) \rightarrow xy = yx$. By compactness, we can replace q by some finite part $q_0 \subset q$. As p is a non-realized type, the definable set $q_0(x)$ is infinite and has the required property. \square

Proposition 2.26. *Let G be a definable group, and assume that there is an infinite subset $A \subset G$ such that any two elements of A commute, then there is an infinite abelian definable subgroup of G .*

Proof. Let \mathcal{D} be an ultrafilter on A and consider the average type $p_{\mathcal{D}}$ of that ultrafilter. Then $p_{\mathcal{D}}$ satisfies the conditions of the previous lemma. \square

Note that if G is torsion free, then the proposition is easy (and does not require NIP): take $H = C(C(a))$ for any $a \in G \setminus \{1\}$.

2.2.2 Eventual types

In a stable theory, an invariant type p is the average of any one of its Morley sequences. This is no longer true in general for NIP theories. We show now

that we can nevertheless recover an A -invariant type from its Morley sequence over A .

Proposition 2.27. *Let $p, q \in S_x(\mathcal{U})$ be A -invariant types. If $p^{(\omega)}|_A = q^{(\omega)}|_A$, then $p = q$.*

Proof. By assumption, if I is a Morley sequence of p over A (indexed by any linear order) then it is also a Morley sequence of q over A . Let $\phi(x; d) \in L(\mathcal{U})$ be any formula and let $I = (a_i : i \in \mathcal{I})$ be a Morley sequence of p over A such that $\text{alt}(\phi(x; d), I)$ is maximal. Let $\eta \in \{0, 1\}$ be such that $\lim(I) \vdash \phi(x; d)^\eta$. Now let $a \models p|_{Id}$. Then the sequence $I + (a)$ is also a Morley sequence of p over A . By maximality of I , we must have $\text{alt}(\phi(x; d), I) = \text{alt}(\phi(x; d), I + (a))$. This implies that $a \models \phi(x; d)^\eta$. Therefore $p \vdash \phi(x; d)^\eta$ and as the roles of p and q are interchangeable, we also have $q \vdash \phi(x; d)^\eta$. Thus $p = q$. \square

Definition 2.28. Let $A \subset \mathcal{U}$ and let I be an A -indiscernible sequence. We say that I is *based* on A if for all A -indiscernible sequences I_1, I_2 of same EM-type as I over A , there is some tuple a such that $I_1 + (a)$ and $I_2 + (a)$ are both indiscernible over A .

Let I be any indiscernible sequence based on A . Let $\phi(x; d) \in L(\mathcal{U})$ be any formula. Let $I_1 \equiv_A^{EM} I$ and $I_2 \equiv_A^{EM} I$ such that $\text{alt}(\phi(x; d), I_i)$ is maximal, for $i \in \{1, 2\}$. Let η_i , for $i = 1, 2$ be such that $\lim(I_i) \vdash \phi(x; d)^{\eta_i}$. As I is based on A , there is some tuple a such that both $I_1 + (a)$ and $I_2 + (a)$ are indiscernible over A . By maximality of I_1 , we have $\text{alt}(\phi(x; d), I_1 + (a)) = \text{alt}(\phi(x; d), I_1)$ and $a \models \phi(x; d)^{\eta_1}$. Similarly we have $a \models \phi(x; d)^{\eta_2}$ therefore $\eta_1 = \eta_2$.

We now define a global type $Ev(I/A)$, called the *eventual type* of the sequence I over A . For this, let $\phi(x; d) \in L(\mathcal{U})$. Pick some $J \equiv_A^{EM} I$ such that $\text{alt}(\phi(x; d), J)$ is maximal, and set $Ev(I/A) \vdash \phi(x; d)$ if and only if $\lim(J) \vdash \phi(x; d)$. By the previous argument, this does not depend on the choice of J . It is clear that this defines a consistent, complete type over \mathcal{U} . Also, $Ev(I/A)$ only depends on the type of I over A (indeed of its EM-type) therefore $Ev(I/A)$ is an A -invariant type.

Another characterization of $Ev(I/A)$ (which explains its name) is the following: for any formula $\phi(x; d) \in L(\mathcal{U})$, we have $Ev(I/A) \vdash \phi(x; d)$ if and only if for every sequence $J \equiv_A^{EM} I$, one can increase J to some A -indiscernible sequence J' with $\lim(J') \vdash \phi(x; d)$.

EXAMPLE 2.29. Let T be DLO, and consider an increasing sequence $I = (a_i : i < \omega)$. Let $b_2 < b_1$ be two points greater than all the a_i 's. Set $A_1 = \{b_1\}$ and

$A_2 = \{b_1, b_2\}$, then I is indiscernible both over A_1 and A_2 . The type $Ev(I/A_1)$ is the definable type “ b_1^- ” whereas $Ev(I/A_2)$ is the type “ b_2^- ”. This shows that changing the parameter set A may modify the eventual type $Ev(I/A)$.

Let I be an indiscernible sequence based on A . We show that I is a Morley sequence of $Ev(I/A)$. Without loss, we may assume that $I = (a_i : i < \omega)$. It is enough to show that for each $n < \omega$, we have $a_n \models Ev(I/A) \upharpoonright Aa_{<n}$. We know that we can increase I to an A -indiscernible sequence J such that $\lim(J) \upharpoonright Aa_{<n} = Ev(I/A) \upharpoonright Aa_{<n}$. By indiscernability of J , $a_n \models \lim(J) \upharpoonright Aa_{<n}$ which gives what we want.

Proposition 2.30. *There is a bijection between EM-classes over A of indiscernible sequences based on A and A -invariant types.*

Proof. From left to right, we map an indiscernible sequence I to $Ev(I/A)$ which only depends on the EM-type of I over A . In the other direction, we map an A -invariant type p to the EM-class of its Morley sequence. \square

Proposition 2.31. *The theory T is NIP if and only if for all $M \models T$, for all $p \in S(M)$, p has at most $2^{|M|+|T|}$ global coheirs.*

Proof. We show left to right. Assume that T is NIP and let $M \models T$, then by Proposition 2.27 an M -invariant type is determined by $p^{(\omega)}|_M$. There are at most $2^{|M|+|T|}$ values for that type, so there are at most $2^{|M|+|T|}$ global M -invariant types. The result follows.

Conversely, assume that T has IP. Then there is a set $\{a_i : i < \lambda\}$ of finite tuples, for $\lambda = |T|^+$, and a formula $\phi(x; y)$ such that for any $A \subseteq \lambda$, we can find some b_A such that: $\phi(a_i; b_A) \iff i \in A$. Let M be a model of size λ containing all the a_i 's. For \mathcal{D} an ultrafilter over λ , we define $p_{\mathcal{D}}$ as the average of $\text{tp}(a_i/\mathcal{U})$ along \mathcal{D} . Then we have $p_{\mathcal{D}} \vdash \phi(x; b_A) \iff A \in \mathcal{D}$. This shows that for $\mathcal{D} \neq \mathcal{D}'$, the two types $p_{\mathcal{D}}$ and $p_{\mathcal{D}'}$ are distinct. Furthermore, $p_{\mathcal{D}}$ is finitely satisfiable in M . As there are 2^{2^λ} pairwise distinct ultrafilters on λ , we have obtained 2^{2^λ} global types finitely satisfiable in M . Since there are only 2^λ types of the same arity over M , there is at least one such type which has 2^{2^λ} global coheirs. \square

2.2.3 Generically stable types

Lemma 2.32. *Let $I = (a_i : i \in \mathcal{I})$ be a totally indiscernible sequence, and $\phi(x; b) \in L(\mathcal{U})$ a formula ($|x| = |a_i|$), then the set $\{i \in \mathcal{I} : \models \phi(a_i; b)\}$ is finite or cofinite in \mathcal{I} .*

Proof. Otherwise, we can build a sequence $(i_k : k < \omega)$ of pairwise distinct members of \mathcal{I} such that $i_k \in \mathcal{I}$ if and only if k is even. Then the sequence $(a_{i_k} : k < \omega)$ is indiscernible and the formula $\phi(a_{i_k}; b)$ holds if and only if k is even. This contradicts NIP. \square

Note that it follows from the proof that the cardinality of $\{i \in \mathcal{I} : \models \phi(a_i; b)\}$ or of its complement is bounded by $\text{alt}(\phi(x; y))/2$.

Theorem 2.33. *Let p be a global A -invariant type. Then the following are equivalent:*

- (i) *if $I \models p^{(\omega)}|_A$, then $p = \lim(I)$;*
- (ii) *p is definable and finitely satisfiable;*
- (iii) *$p_x \otimes p_y = p_y \otimes p_x$;*
- (iv) *any Morley sequence of p is totally indiscernible.*

Proof. (i) \Rightarrow (ii): Assume (i). Then p is finitely satisfiable in I . Also, let $\phi(x; y) \in L$ and let $b \in \mathcal{U}$. Then $p \vdash \phi(x; b)$ if and only if there is some sequence $(a_k : k < \omega) \models p^{(\omega)}|_A$ such that $\phi(a_k; b)$ holds for all k . Therefore $\{b \in \mathcal{U} : p \vdash \phi(x; b)\}$ is a type-definable set. The same holds for $\neg\phi$ instead of ϕ , so the complement of that set is also type-definable. It follows that it is definable. Therefore p is a definable type.

(ii) \Rightarrow (iii): Follows from Lemma 2.23.

(iii) \Rightarrow (iv): By associativity of \otimes , we have that for every n and $\sigma \in \text{Sym}(n)$, $p(x_{n-1}) \otimes \dots \otimes p(x_0) = p(x_{\sigma(n-1)}) \otimes \dots \otimes p(x_{\sigma(0)})$. It follows that the Morley sequence of p is totally indiscernible.

(iv) \Rightarrow (i): Let $\phi(x; b) \in L(\mathcal{U})$ and $I \models p^{(\omega)}|_A$. Let $J \models p^{(\omega)}|_{A \cup b}$. Then $I + J$ is a totally indiscernible sequence. Let $\epsilon \in \{0, 1\}$ be such that $p \vdash \phi(x; b)^\epsilon$. Then all points of J satisfy $\phi(x; b)^\epsilon$. Therefore by Lemma 2.32, at most a finite number of points of I satisfy $\neg\phi(x; b)^\epsilon$. It follows that $\lim(I) \vdash \phi(x; b)^\epsilon$. \square

Definition 2.34. An invariant type satisfying the equivalent conditions of Theorem 2.33 is called *generically stable*.

EXAMPLE 2.35. 1. Let $(M; \leq, \wedge)$ be a model of the theory T_{dt} of dense trees as defined in Appendix A. Let $c \in M$. We define the generic type of the closed cone centered in c as the global type $p_c(x)$ satisfying $x > c$ and $\neg(x \leq b)$ for any $b \in \mathcal{U}$, $b > c$. This is a complete, c -invariant type. Its Morley sequence $(b_k : k < \omega)$ is a sequence of pairwise incomparable elements, such that $b_k \wedge b_l = c$ for any $k \neq l$. Using quantifier elimination, we see that it is a totally indiscernible sequence. Therefore p_c is generically stable.

2. Consider the language $L = \{R_n(x, y) : n < \omega\}$ and the L -structure M whose universe is \mathbb{Q} and such that $M \models R_n(x, y)$ if and only if $x < y$ and $|x - y| < n$. This is sometimes called a “local order”.

Let $p(x)$ be the global \emptyset -invariant type satisfying $\neg R_n(a, x) \wedge \neg R_n(x, a)$ for all $a \in \mathcal{U}$. Then this type is generically stable.

Remark 2.36. If p is generically stable, then it is definable and finitely satisfiable over any one of its Morley sequences. (Because if I is a Morley sequence of p , then $p = \lim(I)$ is finitely satisfiable in I , furthermore it is I invariant, so it is definable over I). We can write a definition explicitly. Let $\phi(x; y)$ be a formula, and let $N = \text{alt}(\phi)$. If $(a_k : k < \omega)$ is a Morley sequence of p , and $b \in \mathcal{U}$, then

$$p \vdash \phi(x; b) \iff \bigvee_{\substack{A \subset N+1 \\ |A| > N/2}} \bigwedge_{k \in A} \phi(a_k; b).$$

Notice in particular that the form of the formula giving the ϕ -definition depends only on ϕ .

Proposition 2.37. *Let p be generically stable and q any invariant type. Then $p_x \otimes q_y = q_y \otimes p_x$.*

Proof. Assume for a contradiction that for some formula $\phi(x; y) \in L(\mathcal{U})$ we have $p_x \otimes q_y \vdash \phi(x; y)$ and $q_y \otimes p_x \vdash \neg \phi(x; y)$. Let $(a_k : k < \omega) \models p^{(\omega)}$, $b \models q|_{\mathcal{U}_{<\omega}}$ and $(a_k : \omega \leq k < \omega 2) \models p^{(\omega)}|_{\mathcal{U}_{<\omega}} b$. Then for $k < \omega$, $\neg \phi(a_k; b)$ holds and for $k \geq \omega$, we have $\phi(a_k; b)$. As the sequence $(a_k : k < \omega 2)$ is totally indiscernible, this contradicts Lemma 2.32. \square

Lemma 2.38. *Let p be generically stable, then for any integer n , $p^{(n)}$ is also generically stable.*

Proof. One checks easily that a product of definable types is definable, and a product of finitely satisfiable types is finitely satisfiable. Therefore a product of generically stable types is again generically stable. \square

Proposition 2.39. *Let p be a generically stable, A -invariant type. Then p is the unique A -invariant extension of $p|_A$.*

Proof. Notice that if q is any A -invariant type and $r, s \in S(\mathcal{U})$ are such that $r|_A = s|_A$, then $(q \otimes r)|_A = (q \otimes s)|_A$.

Let q be any A -invariant extension of $p|_A$. We will show by induction that $q^{(n)}|_A = p^{(n)}|_A$. It follows by Proposition 2.27 that $p = q$. The case $n = 1$ is the hypothesis on q . Assume that $q^{(n)}|_A = p^{(n)}|_A$. We have:

$$\begin{aligned}
 q^{(n+1)}(x_1, \dots, x_{n+1})|_A &= (q(x_{n+1}) \otimes q^{(n)}(x_1, \dots, x_n))|_A \\
 &= (q(x_{n+1}) \otimes p^{(n)}(x_1, \dots, x_n))|_A \\
 &= (p^{(n)}(x_1, \dots, x_n) \otimes q(x_{n+1}))|_A \\
 &= (p^{(n)}(x_1, \dots, x_n) \otimes p(x_{n+1}))|_A \\
 &= p^{(n+1)}(x_1, \dots, x_{n+1})|_A.
 \end{aligned}$$

□

Historical Remarks

The definition of NIP and most results in Section 2.1 are due to Shelah and appeared first in [78] (see also [81][II.4]). Lemma 2.5 comes from Poizat's [70]. The characterization 2.31 also appears in that paper, but the proof we give here is from [80]. Eventual types are defined in [71], although the terminology was introduced by Adler in [4]. Proposition 2.14 was first proved by Shelah using the $\text{ded}(\lambda)$ -characterization and the fact that the statement is absolute, see the discussion in [72][Chapter 12]. The proof we give here is very close to Poizat's approach in [72][12.18], and was first published by Adler in [4].

The Baldwin-Saxl theorem is from [12]. See also [69]. Proposition 2.26 was proved by Shelah in [83].

Generically stable types are first defined by Shelah as 'stable types' in [82]. They are renamed and studied systematically by Hrushovski and Pillay in [47] and independently by Usvyatsov in [89]. This notion was extended outside of the NIP setting by Pillay and Tanovic in [68].

Chapter 3

Honest definitions and applications

A fundamental characteristic property of stable theories is *definability of types*, namely the property that if $A \subset \mathcal{U}$ is any subset (big or small), and $\phi(x; b) \in L(\mathcal{U})$, a formula, then the set $\phi(A; b)$ coincides with the trace on A of some A -definable set: there is $\psi(x; c) \in L(A)$ such that $\phi(A; b) = \psi(A; c)$. In other words, we can internalize the parameters of ϕ inside A (up to changing the formula). In this chapter, we show that a weak form of definability of types holds in NIP theories. We do not manage to find a definition of $\phi(A; b)$ with parameters inside A , but we do with parameters in some elementary extension A' of A . Furthermore, the definition satisfies a property called *honesty* which says that, on A' , the new formula lies inside the original one.

Associated with definability of types are so-called *reflexion principles* stating that if some (e.g. type-definable) set A is internally simple in some sense, then it is also externally simple: its interactions with the rest of the structure cannot be too complicated. We obtain some statements of this kind using honest definitions. Note that the characteristic property 2.5 of NIP is an example of such a phenomenon. It says that if a sequence I is indiscernible (internal simplicity), then the intersection of I with some unary \mathcal{U} -definable set is a finite union of convex subsets of I (external simplicity). We will in fact extend this result to definable sets of higher arity (*shrinking of indiscernibles*).

Throughout this chapter, we assume that T is NIP.

3.1 Honest definitions

Pairs and definable types

Let M be an L -structure and let $A \subseteq M$. We define the language $L_{\mathbf{P}} = L \cup \{\mathbf{P}(x)\}$, where $\mathbf{P}(x)$ is a new unary relation symbol. We will sometimes write $x \in \mathbf{P}$ instead of $\mathbf{P}(x)$.

The *pair* (M, A) is the $L_{\mathbf{P}}$ -structure whose L -reduct is M , and such that $\mathbf{P}(M) = A$.

Let $A \subseteq M$ be any subset and let $b \in M^k$ and $\phi(x; y) \in L$. We say that the ϕ -type of d over A is definable if there is a formula $\psi(x; z)$ and some tuple $d \in A^{k'}$ such that $\phi(A; b) = \psi(A; d)$. Note that this is expressible in the pair (M, A) by the $L_{\mathbf{P}}$ -formula $\forall x \in \mathbf{P} \phi(x; b) \leftrightarrow \psi(x; d)$.

In a stable theory, any type over any set is definable (see Appendix B). This is not true in an unstable theory, however we will see now that a weaker property holds in NIP theories.

Honest definitions

Theorem 3.1. *Let $M \models T$, $A \subseteq M$, $\phi(x; y) \in L$ a formula and $b \in M^{|y|}$ a tuple. Then there is an elementary extension $(M, A) \prec (M', A')$, a formula $\psi(x; z) \in L$ and a tuple d of elements of A' such that*

$$\phi(A; b) \subseteq \psi(A'; d) \subseteq \phi(A'; b).$$

Remark 3.2. Note that the conclusion is equivalent to saying that there is a formula $\psi(x; z)$ such that for any finite $A_0 \subseteq \phi(A; b)$, there is a tuple d of elements of A such that $A_0 \subseteq \psi(A; d) \subseteq \phi(A; b)$.

EXAMPLE 3.3. Take for example T to be DLO, $M = (\mathbb{R}, <)$ and $A = \mathbb{Q}$. Let $\phi(x; b)$ be $x \leq \pi$ and let $\psi(x; z) = x \leq z$. We see that for any finite $A_0 \subset \phi(A; b)$, there is some point $d \in A$ such that $A_0 \subseteq \psi(A; d) \subseteq \phi(A; b)$. Simply take d to be a rational smaller than π , but greater than all the elements of A_0 .

Proof of Theorem 3.1. Let $S_A \subset S_x(\mathcal{U})$ be the set of global types (in the variable x), finitely satisfiable in A . It is a closed set of $S_x(\mathcal{U})$, and therefore compact. Let $(M', A') \succ (M, A)$ be a $|M|^+$ -saturated elementary extension. Let $p \in S_A$. We try to build a sequence $(a_i : i < \omega)$ such that for all i we have

- $a_i \in A$;
- $a_i \models p \upharpoonright Aa_{<i}$;
- $\neg(\phi(a_i; d) \leftrightarrow \phi(a_{i+1}; d))$.

If we succeed, then the sequence $(a_i : i < \omega)$ is a Morley sequence of p over A and as such it is indiscernible. Then the third condition implies that ϕ has infinite alternation rank, contradicting NIP. We conclude that the construction must stop at some finite stage. So assume we have built $(a_i : i < n)$, and we cannot find a point a_n . Let $\epsilon_p \in \{0, 1\}$ be such that $p \vdash \phi(x; b)^{\epsilon_p}$. We first argue that $\phi(a_{n-1}; b)^{\epsilon_p}$ holds: the type p is finitely satisfiable in A , therefore for every subset $B \subset M'$ of size $\leq |M|$, the $L_{\mathbf{P}}$ -type $p|_B(x) \cup \{\mathbf{P}(x)\}$ is finitely satisfiable in A and thus realized in A' . Taking B to contain $da_{<n}$ we see that if $\models \neg\phi(a_{n-1}; b)^{\epsilon_p}$, we could find a_n as required.

By compactness, there is a formula $\theta_p(x) \in L(Aa_{<n})$ such that $p \vdash \theta_p(x)$ and $(M', A') \models (\forall x \in \mathbf{P})\theta_p(x) \rightarrow \phi(x; d)^{\epsilon_p}$. As S_A is compact, we can find a finite set $S_0 \subset S_A$ such that $\bigcup_{p \in S_0} \theta_p(x)$ covers S_A . Set

$$\psi(x) = \bigvee_{\substack{p \in S_0 \\ \epsilon_p = 1}} \theta_p(x).$$

Write $\psi(x) = \psi(x; d)$ where $d \in A'$ are the parameters appearing in ψ . We show that ψ has the required properties. First, for all type $p \in S_A$ such that $p \vdash \phi(x; b)$, we have $p \vdash \psi(x; d')$. In particular, this is true for realized types. It follows that $\phi(A; b) \subseteq \psi(A'; d)$. Furthermore, we have $(M', A') \models (\forall x \in \mathbf{P})\psi(x; d) \rightarrow \phi(x; d)$, therefore $\psi(A', d) \subseteq \phi(A, b)$. \square

Definition 3.4. We say that the formula $\psi(x; d)$ in the previous theorem is an *honest definition* of $\phi(x; b)$ over A .

The following corollary is simply the weakening of the theorem obtained by removing in the conclusion the “honesty” hypothesis $\psi(A'; d) \subseteq \phi(A'; b)$. We state it separately because it might seem more natural and because it is sufficient for some applications. We will refer to it as *weak stable-embeddedness* of the set A .

Corollary 3.5 (Weak stable-embeddedness). *Let $M \models T$, $A \subset M$, $\phi(x; y) \in L$ and $b \in M^{|y|}$. Then there is an elementary extension $(M, A) \prec (M', A')$, a formula $\psi(x; z) \in L$ and a tuple $d \in A'$ such that*

$$\phi(A; b) = \psi(A; d).$$

Corollary 3.6. *Let $M \models T$ and $A \subseteq M$. Let $b \in M$ be a finite tuple. Let $(M, A) \prec (M', A')$ be an $|M|^+$ -saturated extension. Then there is $A_0 \subset A'$ of size at most $|T|$ such that for any two tuples a and a' from A we have:*

$$a \equiv_{A_0} a' \implies a \equiv_b a'.$$

Proof. For any formula $\phi(x; y)$, $|y| = |b|$, let $\psi_\phi(x; d_\phi)$ be an honest definition of $\phi(x; b)$ over A where $d_\phi \in A'$. Take A_0 to contain the union of the parameters d_ϕ for ϕ ranging over all formulas. \square

Examples

Let $A \subseteq M$ be any subset and let $A_{\text{ind}(\emptyset)}$ be the structure in the language $L_{A, \emptyset} = \{R_{\phi(\bar{x})}(\bar{x}) : \phi(\bar{x}) \in L(\emptyset)\}$ whose universe is A and where each $R_{\phi(\bar{x})}$ is interpreted the obvious way: for every $\bar{a} \in A$, $A_{\text{ind}(\emptyset)} \models R_{\phi(\bar{a})} \iff M \models \phi(\bar{a})$.

Proposition 3.7. *Let $A \subseteq M$. Assume that in $A_{\text{ind}(\emptyset)}$, the quantifier-free formulas are stable. Then A is stably-embedded in M .*

Proof. By weak stable-embeddedness of A (Corollary 3.5), there is $(M', A') \succ (M, A)$ and $\psi(x; d) \in L(A')$ such that $\phi(A; b) = \psi(A; d)$. Note that we have $A_{\text{ind}(\emptyset)} \prec A'_{\text{ind}(\emptyset)}$ and $\psi(A; d) = R_{\psi(x; z)}(A; d)$. We work inside the structure $A'_{\text{ind}(\emptyset)}$. As the formula $R_{\psi(x; z)}(x; z)$ is stable, the set $R_{\psi(x; z)}(A; d)$ is definable by some R_ψ -formula (in particular quantifier-free) with parameters in A (see Appendix B). This translates into some formula $\theta(x; c) \in L(A)$ such that $\theta(A; c) = \phi(A; d)$ as required. \square

EXERCISE 3.8. *Let $D \subseteq M$ be a \emptyset -definable set. Let $\phi(x, y)$ be a \emptyset -definable formula which defines a total order $<$ on D . Assume that $D_{\text{ind}(\emptyset)}$ is o-minimal when equipped with that ordering. Then any subset of D definable in M is a finite union of $<$ -convex sets.*

3.2 Naming a submodel

Let M be a model of T . A subset $D \subseteq M^k$ of the form $\phi(M; b)$ for some $\phi(x; b) \in L(\mathcal{U})$ is called an *externally definable set* of M and the formula $\phi(x; b)$ is called an *external definition* of D .

Proposition 3.9. *Let $D \subseteq M^k$ be an externally definable set. Then there is an external definition $\phi(x; b) \in L(\mathcal{U})$ of D with the following property:*

for every formula $\theta(x; a) \in L(M)$, $|x| = k$, such that $D \subseteq \theta(M; a)$, we have $\mathcal{U} \models \phi(x; b) \rightarrow \theta(x; a)$.

Proof. Let $M \prec N$, with N an $|M|^+$ -saturated model. Let $\phi_0(x; b_0) \in L(N)$ be an external definition of D . Theorem 3.1 applied to the pair (N, M) and the formula $\phi_0(x; b_0)$ yields an elementary extension $(N, M) \prec (N', M')$ and a formula $\phi(x; b) \in L(M')$ such that $D \subseteq \phi(x; b) \subseteq \phi_0(x; b_0)$. Let $\theta(x; a) \in L(M)$ be a formula such that $D \subseteq \theta(M; a)$. Then $(N, M) \models (\forall x \in \mathbf{P}) \phi_0(x; b_0) \rightarrow \theta(x; a)$. Therefore the same sentence holds in the pair (N', M') . We conclude that $\phi(M'; b) \subseteq \phi_0(M'; b_0) \subseteq \theta(M'; a)$. As M' is a model, and $b \in M'$, we have $M' \models \phi(x; b) \rightarrow \theta(x; a)$, as required. \square

If M is a model of T we define the *Shelah expansion* of M , denoted M^{Sh} as follows. The language of M^{Sh} is $L_{Sh}(M)$ containing, for each integer k and each externally definable $D \subseteq M^k$, a k -ary predicate $R_D(\bar{x})$. The universe of M^{Sh} is M and the predicates are interpreted in the canonical way. Note that in particular M^{Sh} is an expansion of M and contains a constant for every element of M .

Proposition 3.10. *The structure M^{Sh} admits elimination of quantifiers.*

Proof. We have to show that the projection of an externally definable set is again externally definable. Let $D \subseteq M^{k_1+k_2}$ be externally definable. Let $\phi(x_1, x_2; b) \in L(\mathcal{U})$ be an external definition of D given by Proposition 3.9 (where $|x_1| = k_1$ and $|x_2| = k_2$). Let π denote the projection $M^{k_1+k_2} \rightarrow M^{k_1}$ and set $\psi(x_1; b) = (\exists x_2) \phi(x_1, x_2; b)$. We claim that $\pi(D) = \psi(M; b)$.

It is clear that $\pi(D) \subseteq \psi(M; b)$. To show the other inclusion, take some $a \in M^{k_1} \setminus \pi(D)$. Let $\zeta(x_1, x_2) \equiv x_1 \neq a$. Then $D \subseteq \zeta(M)$. By hypothesis on $\phi(x_1, x_2; b)$, we have $\phi(x_1, x_2; b) \rightarrow \zeta(x_1, x_2)$, therefore $\psi(x_1; b) \rightarrow x_1 \neq a$. It follows that $a \notin \psi(M; b)$ as required. \square

Corollary 3.11. *The structure M^{Sh} is NIP.*

Proof. This follows from Proposition 3.10, since the formula $R_D(\bar{x}, \bar{y})$ cannot have alternation rank greater than that of $\phi(\bar{x}, \bar{y})$ where $\phi(\bar{x}, \bar{y}) \in L(\mathcal{U})$ is an external definition of D . \square

Remark 3.12. There are structures with IP which satisfy Proposition 3.10. In fact, there are structures M in which every subset of M^k for all k is externally definable. This is the case for example if M a model of arithmetic.

EXERCISE 3.13. Show that there is a canonical bijection between types over \emptyset in the structure M^{Sh} and L -types over \mathcal{U} , finitely satisfiable in M .

EXERCISE 3.14. Let $M \models T$, $\phi(x, y; d) \in L(\mathcal{U})$. Assume that $\phi(M; b)$ is the graph of a function. Then there is a formula $\psi(x, y; d) \in L(\mathcal{U})$ such that $\phi(M; b) = \psi(M; d)$ and $\mathcal{U} \models (\forall x)(\exists^{\leq 1} y)\psi(x, y; d)$ (so $\psi(x, y; d)$ is the graph of a partial function).

3.3 Shrinking of indiscernible sequences

We are interested now in the case where $A = I$ is an indiscernible sequence. We study the trace on I^k of definable sets in k variables. The case $k = 1$ has already been dealt with in Chapter 2: the trace on I of a definable set is a finite union of $<_I$ -convex sets.

We will give several statements. Later results generalize previous ones, so there is some redundancy.

Definition 3.15. Let $(\mathcal{I}, <)$ be a linear order. A *convex equivalence relation* \sim on \mathcal{I} is an equivalence relation all of whose classes are convex sets. The equivalence relation \sim is said to be *finite* if it has finitely many classes.

If \sim is a convex equivalence relation on $(\mathcal{I}, <)$, we extend \sim to cartesian powers of \mathcal{I} in the following way: if $\bar{i} = (i_1, \dots, i_n), \bar{j} = (j_1, \dots, j_n) \in \mathcal{I}^n$, then we set $\bar{i} \sim \bar{j}$ if:

- for all $k \leq n$, $i_k \sim j_k$;
- for all $k, k' \leq n$, we have $i_k < i_{k'} \iff j_k < j_{k'}$ and $i_k = i_{k'} \iff j_k = j_{k'}$.

If $c \in \mathcal{I}$, we define \sim_c as the equivalence relation on \mathcal{I} defined by: $i \sim_c j$ if either $i = j = c$, or $(i < c \text{ and } j < c)$, or $(i > c \text{ and } j > c)$. If $\bar{c} \subset \mathcal{I}$ is a family (or a set) of elements, we define $\sim_{\bar{c}}$ as the intersection of the relations \sim_c for c ranging in \bar{c} .

Proposition 3.16. Let $I = (a_i : i \in \mathcal{I})$ be an indiscernible sequence. Let $\phi(x_1, \dots, x_n; b) \in L(\mathcal{U})$ be a formula. Then there is a finite convex equivalence relation \sim on \mathcal{I} such that for all $\bar{i} \in \mathcal{I}^n, \bar{j} \in \mathcal{I}^n$, we have

$$\bar{i} \sim \bar{j} \Rightarrow \phi(a_{\bar{i}}; b) \leftrightarrow \phi(a_{\bar{j}}; b).$$

Proof. Without loss of generality, and for simplicity of notations, we assume that I is a sequence of singletons (since we can work in M^{eq} for example).

First assume that I is totally indiscernible. The structure $I_{ind(\emptyset)}$ is an expansion by definitions (naming constants) of I with pure equality. Therefore by Proposition 3.7, the set I is stably embedded. Let $\psi(x_1, \dots, x_n; m)$, with $m \in I$ be such that $\psi(I; m) = \phi(I; m)$. Without loss, $\psi(I; m)$ is a formula in the language of pure equality. Then the relation \sim_m has the required property.

Assume now that I is not totally indiscernible. Let M be some $|I|^+$ -saturated model containing I . There is some formula $\theta(x, y) \in L(M)$ which orders I (namely such that $M \models \theta(a_i, a_j) \iff i < j$). Consider the pair (M, I) . By Corollary 3.5, there is an elementary extension $(M, I) \prec (M', I')$ and a formula $\psi(x_1, \dots, x_n; d) \in L(I')$ such that $\phi(I; b) = \psi(I'; d)$. Note that the formula $\theta(x, y)$ defines a linear order on I' and thus ordered, I' is an indiscernible sequence. Write $I' = (a_i : i \in \mathcal{I}')$ with $\mathcal{I} \subset \mathcal{I}'$. It follows that the formula $\psi(x_1, \dots, x_n; d)$ is equivalent on I' , to some boolean combinations of formulas of the form $x_i = d_k$, $\theta(x_i, d_k)$, $x_i = x_j$, $\theta(x_i, x_j)$ where each d_k is a element of the tuple d . We define \sim on \mathcal{I} as being the restriction of the relation \sim_d of \mathcal{I}' . We see that \sim has the required property. \square

Theorem 3.17. *Let A be a small set of parameters and $I = (a_i : i \in \mathcal{I})$ an indiscernible sequence. Let $\phi(\bar{x}, y; b) \in L(\mathcal{U})$ with $\bar{x} = (x_1, \dots, x_n)$. Then there is a finite convex equivalence relation \sim on \mathcal{I} such that for $\bar{i}, \bar{j} \in \mathcal{I}^n$, we have*

$$\bar{i} \sim \bar{j} \Rightarrow \forall e \in A, \phi(a_{\bar{i}}, e; b) \leftrightarrow \phi(a_{\bar{j}}, e; b).$$

Proof. The proof is similar to that of Proposition 3.16, except that instead of naming the sequence I , we name the product $A \times I$ (as a subset of M^{eq}).

Take M a sufficiently saturated model containing A and I and consider the pair $(M, A \times I)$. By Corollary 3.5, there is an elementary extension $(M, A \times I) \prec (M', A' \times I')$ and a formula $\psi(x_1, \dots, x_n, y; a, m)$, $a \in A$, $m \in I'$ such that $\phi(I \times A; b) = \psi(I \times A; a, m)$. By elementarity, the set I' can be ordered as $I' = (a_i : i \in \mathcal{I}')$, with $\mathcal{I} \subset \mathcal{I}'$, such that it becomes a sequence indiscernible over A' (this is true also if I is totally indiscernible, in which case we can take any order on I'). One then checks that the relation \sim on \mathcal{I} defined as the restriction of \sim_m on \mathcal{I}' has the required property. \square

Remark 3.18. If the indiscernible sequence I is indexed by a complete order \mathcal{I} , then the relation \sim in Theorem 3.1 can be taken to be of the form $\sim_{\bar{c}}$ for some $\bar{c} \in I^n$. (Because any finite convex equivalence relation on \mathcal{I} is of that form).

If furthermore there is a formula $\theta(x, y)$ which orders I and has parameters in I , then I is stably embedded.

(This may fail even if I is not totally indiscernible. Take for example a circular order $(M, R(x, y, z))$, where $R(x, y, z)$ is the betweenness relation and an indiscernible sequence $I = (a_i : i \in \mathbb{Q})$ of points from it. Then the involution $a_k \mapsto a_{-k}$ is elementary, so we cannot define the order on I without extra parameters. This is the only possible obstruction; see [41]).

We say that a convex equivalence relation \sim is *essentially of size κ* if it is the intersection of κ many finite convex equivalence relations. Note that a relation essentially of size κ has at most 2^κ many classes.

Remark 3.19. If \mathcal{I} is a linear order, let $\text{compl}(\mathcal{I})$ be the completion of \mathcal{I} . Let \sim be any convex equivalence relation on \mathcal{I} essentially of size κ . Then there is $\bar{c} \subseteq \mathcal{J}$ of size $\leq \kappa$ such that the restriction of $\sim_{\bar{c}}$ to \mathcal{I} refines \sim and for every $i, j \in \mathcal{I} \setminus \bar{c}$, we have $i \sim j \iff i \sim_{\bar{c}} j$.

Therefore when given a convex equivalence relation on \mathcal{I} which is finite (resp. essentially of size κ), we may assume that it is of the form $\sim_{\bar{c}}$ for some finite $\bar{c} \subseteq \text{compl}(\mathcal{I})$ (resp. \bar{c} of size $\leq \kappa$).

Corollary 3.20. *Let A be a small set of parameters and $I = (a_i : i \in \mathcal{I})$ an indiscernible sequence of finite tuples. Let $b \in \mathcal{U}$ be a finite tuple. Then there is a convex equivalence relation \sim on \mathcal{I} essentially of size $|\mathcal{I}|$ such that for any $\phi(\bar{x}, y; z) \in L$ with $\bar{x} = (x_1, \dots, x_n)$ for any $\bar{i}, \bar{j} \in \mathcal{I}^n$, we have*

$$\bar{i} \sim \bar{j} \implies \forall e \in A, \phi(a_{\bar{i}}, e; b) \leftrightarrow \phi(a_{\bar{j}}, e; b).$$

Note that we recover the claim stated during the proof of Proposition 2.14 that if the cofinality of \mathcal{I} is greater than $|\mathcal{I}|$, then there is an end segment of I which is indiscernible over Ab .

It turns out that this result is also true when the sequence I is composed of tuples of infinite length, although we need to redo the proof in that case.

Theorem 3.21. *Let A be a small set of parameters and $I = (\bar{a}_i : i \in \mathcal{I})$ an A -indiscernible sequence of tuples of arbitrary length. Let $b \in \mathcal{U}$ be a finite tuple. Then there is a convex equivalence relation \sim on \mathcal{I} essentially of size $|\mathcal{I}|$ such that for any $\phi(\bar{x}, y; z) \in L$ with $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$ for any $\bar{i}, \bar{j} \in \mathcal{I}^n$, we have*

$$\bar{i} \sim \bar{j} \implies \forall e \in A, \phi(\bar{a}_{\bar{i}}, e; b) \leftrightarrow \phi(\bar{a}_{\bar{j}}, e; b).$$

Proof. Let M be a sufficiently saturated model containing A , I and b . Write $\bar{a}_i = (a_i^j : j < \alpha)$. We consider two new unary predicate $\mathbf{P}(x)$ and $\mathbf{A}(x)$ and three new binary predicates $\mathbf{E}(x, y)$, $\mathbf{F}(x, y)$ and $\mathbf{R}(x, y)$. Let $L' = L \cup \{\mathbf{P}, \mathbf{A}, \mathbf{E}, \mathbf{F}, \mathbf{R}\}$ and expand M into an L' -structure (M, P, A, E, F, R) where $P = \{a_i^j : j < \alpha, i \in \mathcal{I}\} \cup A$, $E = \{(a_i^j, a_i^{j'}) : i \in \mathcal{I}, j, j' < \alpha\}$, $F = \{(a_i^j, a_{i'}^{j'}) : i, i' \in \mathcal{I}, j < \alpha\}$ and $R = \{(a_i^j, a_{i'}^{j'}) : i < i' \in \mathcal{I}, j, j' < \alpha\}$.

Let $(M, P, A, E, F, R) \prec (M', P', A', E', F', R')$ be some $|M|^+$ -saturated elementary extension. Using the extra structure given by E , F and R , we can write $P' = A' \cup \{a_i^j : i \in \mathcal{I}', j < \beta\}$ for some $\beta \geq \alpha$ and $\mathcal{I}' \supseteq \mathcal{I}$ and such that the sequence $((a_i^j)_{j < \beta} : i \in \mathcal{I}')$ is indiscernible over A . By Corollary 3.6, there is some set $P_0 \subset P'$ of size at most $|T|$ such that for any two tuples $a, a' \in P$, we have

$$a \equiv_{P_0} a' \implies a \equiv_b a'.$$

Let $\bar{c} \subset \mathcal{I}'$ be the set consisting of elements $c \in \mathcal{I}'$ for which there is $j < \beta$ such that $a_c^j \in A_0$. The equivalence relation $\sim_{\bar{c}}$ on \mathcal{I}' is a convex equivalence relation essentially of size $|T|$. Let \sim be its restriction to \mathcal{I} . Then \sim has the required property. \square

EXERCISE 3.22 (Critical points). Let $I = (a_i : i \in \mathcal{I})$ be an A -indiscernible sequence of finite tuples. Assume that the order \mathcal{I} is dense complete without end points. Let $b \in \mathcal{U}$ be a finite tuple. Let $\phi(x_1, \dots, x_n, y; z)$ be a formula with $|x_k| = |a_i|$ and $|z| = |b|$. Call an index $i \in \mathcal{I}$ ϕ -critical if there is $k \in \{1, \dots, n\}$, some $e \in A$ and indices $i_1 < \dots < \hat{i}_k < \dots < i_n$ in \mathcal{I} such that for any $i_* < i < i_{**}$ we can find $i_* < i_k, i'_k < i_{**}$ such that:

$$\models \phi(a_{i_1}, \dots, a_{i_k}, \dots, a_{i_n}, e; b) \wedge \neg \phi(a_{i_1}, \dots, a_{i'_k}, \dots, a_{i_n}, e; b).$$

Let \bar{c} denote the set of ϕ -critical points and define $\sim_{\bar{c}}$ as usual.

1. For any two finite tuples $\bar{i}, \bar{j} \in \mathcal{I}^n$ and any $e \in A$, we have

$$(\Delta) \quad \bar{i} \sim_{\bar{c}} \bar{j} \implies \phi(a_{\bar{i}}, e; b) \leftrightarrow \phi(a_{\bar{j}}, e; b).$$

2. For any convex equivalence relation \sim satisfying (Δ) , if $i, j \in \mathcal{I}$ are not critical points, then $i \sim j \implies i \sim_{\bar{c}} j$.
3. There are finitely many ϕ -critical points.

Note that in the previous exercise the relation $\sim_{\bar{c}}$ is definable in the pair (M, I) and each ϕ -critical point is in $dcl(Ab)$ read in that structure.

Historical Remarks

Honest definitions were introduced by Chernikov and Simon in [23]. Corollary 3.5 (weak stable-embeddedness) was proved previously by Guingona in [32] using a different method. Proposition 3.10 (Shelah's expansion) was proved first in the o-minimal case by Baisalov and Poizat in [11] and in full generality by Shelah in [83]. The results on shrinking of indiscernibles were first obtained by Baldwin and Benedikt in [13], without the parameter set A , and generalized by Shelah in [83]. The proofs we give here are new.

Chapter 4

Strong dependence and dp-ranks

We introduce the notion of dp-rank, which is similar to *weight* in stable theories. It is a measure of the complexity of a type in an NIP theory. A theory is *strongly-dependent* if all types have $\text{dp-rank} < \aleph_0$. In such theories, shrinking of indiscernibles can be done at once for all formulas, using only finitely many cuts.

We assume NIP throughout this chapter.

4.1 Mutually indiscernible sequences

Definition 4.1. Let $(I_t : t \in X)$ be a family of sequences and A a set of parameters. We say that the sequences $\{I_t : t \in X\}$ are *mutually indiscernible* over A if for each $t \in X$, the sequence I_t is indiscernible over $A \cup \{I_l : l \in X \setminus \{t\}\}$.

EXAMPLE 4.2. Let $I = (a_i : i \in \mathcal{I})$ be an A -indiscernible sequence. Let $(\mathcal{I}_t : t \in X)$ be a family of pairwise disjoint convex subsets of \mathcal{I} . For $t \in X$, let $I_t = (a_i : i \in \mathcal{I}_t)$. Then the sequences $\{I_t : t \in X\}$ are mutually indiscernible over $A \cup \{a_i : i \in \mathcal{I} \setminus \bigcup_{t \in X} \mathcal{I}_t\}$.

EXAMPLE 4.3. Let $(p_t : t < \alpha)$ be a family of global A -invariant types. Build a family $(I_t : t < \alpha)$ of sequences such that for each $t \in X$, I_t is a Morley sequence of p_t over $A \cup \{I_l : l < t\}$. Then the sequences $\{I_t : t \in X\}$ are mutually indiscernible over A .

[*Proof:* Observe that if I is indiscernible over A , p is a global A -invariant type and $a \models p \upharpoonright AI$, then I is indiscernible over Aa . In particular this holds

if a realizes a Morley sequence of p over AI . It is immediate by invariance of p . More generally, if the sequences $\{I_t : t < \beta\}$ are mutually indiscernible over A , then they remain so over Aa where $a \models p \upharpoonright A \cup \{I_t : t < \beta\}$. Thus the result follows by induction.]

EXAMPLE 4.4. Let $(p_t : t < \alpha)$ be a family of global A -invariant types such that $p_t(x) \otimes p_s(y) = p_s(y) \otimes p_t(x)$ for any $t, s < \alpha$, $t \neq s$. Build an array $(a_t^k : t < \alpha, k < \omega)$ such that $a_t^k \models p_t \upharpoonright A \cup \{a_l^s : (l < k) \text{ or } (l = k \text{ and } s < t)\}$. For $t < \alpha$, define $I_t = (a_k^t : k < \omega)$. Note that it is a Morley sequence of p_t over A .

Then the sequences $\{I_t : t \in X\}$ are mutually indiscernible over A .

[*Proof:* We show that for each $t \in X$, I_t is a Morley sequence of p_t over $\{I_s : s \in X \setminus \{t\}\}$. Let $N < \omega$ and pick pairs

$$(k_0, t_0) > \cdots > (k_{n-1}, t_{n-1}) \in \omega \times \alpha,$$

where $\omega \times \alpha$ is ordered lexicographically. Let $c < n$ be minimal such that $t_c = t$ (we assume such a c exists). We have

$$a_{k_0}^{t_0} \wedge \cdots \wedge a_{k_{n-1}}^{t_{n-1}} \models p_{t_0}(x_0) \otimes \cdots \otimes p_{t_{n-1}}(x_{n-1})|_A.$$

As the type p_{t_c} commutes with p_{t_m} for $m < c$, we have

$$a_{k_0}^{t_0} \wedge \cdots \wedge a_{k_{n-1}}^{t_{n-1}} \models p_{t_c}(x_c) \otimes p_{t_0}(x_0) \otimes \cdots \otimes \widehat{p_{t_c}(x_c)} \otimes \cdots \otimes p_{t_{n-1}}(x_{n-1})|_A.$$

Where $\widehat{p_{t_c}(x_c)}$ means that this term is omitted. Therefore $a_{k_c}^{t_c} \models p_{t_c} \upharpoonright A \cup \{a_{k_m}^{t_m} : m \neq c\}$. It follows that the sequence I_t is a Morley sequence of p_t over $A \cup \{I_s : s \neq t\}$ and we conclude as in the previous example.]

EXERCISE 4.5. Let $\{I_t : t \in X\}$ be mutually indiscernible sequences. For $t \in X$, let $p_t = \lim(I_t)$. Then for $t, s \in X$, $t \neq s$, the types p_t and p_s commute (i.e., $p_t(x) \otimes p_s(y) = p_s(y) \otimes p_t(x)$).

Remark 4.6. Let $(I_t : t \in X)$ be a family of sequences, with $I_t = (a_i^t : i \in \mathcal{I}_t)$. Assume that the sequences $(I_t : t \in X)$ are mutually indiscernible over A . Let \mathcal{J} be a linearly ordered set, and for $t \in X$, let $\sigma_t : \mathcal{J} \rightarrow \mathcal{I}$ be an increasing embedding. Then the sequence $((a_{\sigma_t(i)}^t)_{t \in X} : i \in \mathcal{J})$ is indiscernible over A .

Proposition 4.7. Let $(I_t : t \in X)$ be a family of sequences mutually indiscernible over A . Let $b \in \mathcal{U}$ be a finite tuple. Then there is a set $X_b \subset X$ of size at most $|T|$ such that the sequences $\{I_t : t \in X \setminus X_b\}$ are mutually indiscernible over Ab .

Proof. For simplicity assume that the sequences I_t are sequences of finite tuples. (The case of infinite tuples can be taken care of as in the proof of Theorem 3.21). By working in M^{eq} , we may assume that they are sequences of singletons. Write $I_t = (a_i^t : i \in \mathcal{I}_t)$. Let $\mathbf{P}(x)$ and $\mathbf{A}(x)$ be two new unary predicates, and $\mathbf{E}(x, y)$, $\mathbf{R}(x, y)$ be new binary predicates. Let $L' = L \cup \{\mathbf{P}, \mathbf{A}, \mathbf{E}, \mathbf{R}\}$. We define an L' -expansion (M, P, A, E, R) of M by interpreting $\mathbf{P}(x)$ as the set $P = \{a_i^t : t \in X, i \in \mathcal{I}_t\} \cup A$, $\mathbf{A}(x)$ as the set A , $\mathbf{R}(x, y)$ as the set $R = \{(a_i^t, a_{i'}^{t'}) : t \in X, i < i' \in \mathcal{I}_t\}$.

Let $(M, P, A, R) \prec (M', P', A', R')$ be some $|M|^+$ -saturated elementary extension. Using the extra predicates added to the language, we see that there is some $X' \supseteq X$ and sets \mathcal{I}_t for $t \in X'$, with $\mathcal{I}_t \subseteq \mathcal{I}'_t$ when $t \in X$ such that P' can be written as $P' = A' \cup \{a_i^t : t \in X', i \in \mathcal{I}'_t\}$ and such that the sequences $\{I'_t = (a_i^t : i \in \mathcal{I}'_t) : t \in X'\}$ are mutually indiscernible over A' .

By Corollary 3.6, there is some $P_0 \subset P'$ of size at most $|T|$ such that for any finite tuples $a, a' \in P$ we have $a \equiv_{P_0} a' \implies a \equiv_b a'$. Let $X_b \subset X$ be the set of elements $t \in X$ such that some a_i^t belongs to P_0 . Then X_b has size at most $|T|$ and we see that the sequences $\{I_t : t \in X \setminus X_b\}$ are mutually indiscernible over Ab .

□

EXERCISE 4.8. Let $(p_t : t < |T|^+)$ be a family of global A -invariant types such that $p_i(x) \otimes p_j(y) = p_j(y) \otimes p_i(x)$ for all $i \neq j$. Let q be any global invariant type. Then there is $t < |T|^+$ such that $q(x) \otimes p_t(y) = p_t(y) \otimes q(x)$.

[Hint: Build first sequences $(I_t : t < |T|^+)$, each one being a Morley sequence of p_t over the previous ones. Then take b realizing q over those sequences. Finally realize again a family $(J_t : t < |T|^+)$ of Morley sequences of the p_t 's over what has been constructed so far. Notice that the sequences $(I_t + J_t : t < |T|^+)$ are mutually indiscernible and apply Proposition 4.7.]

EXERCISE 4.9. Let $(I_t : t \in X)$ be a family of mutually indiscernible sequences, indexed by the same linear order \mathcal{I} . Write $I_t = (a_i^t : i \in \mathcal{I})$. Define $I = ((a_i^t : t \in X) : i \in \mathcal{I})$. If A is any set such that I is indiscernible over A , then the sequences $\{I_t : t \in X\}$ are mutually indiscernible over A .

EXERCISE 4.10. Let $(I_t : t < \alpha)$ be a family of mutually indiscernible endless sequences. For $t < \alpha$, let $p_t = \lim(I_t)$. Let \mathcal{J} be any linear order. Construct sequences $(J_t : t < \alpha)$, $J_t = (a_i^t : i \in \mathcal{J})$ such that for each $t < \alpha$ and $i \in \mathcal{J}$, $a_i^t \models p_t \upharpoonright \bigcup \{I_s : s < \alpha\} \cup \{a_j^s : s < t \text{ or } (s = t \text{ \& } i < j)\}$. Then the sequences $\{I_t + J_t : t < \alpha\}$ are mutually indiscernible.

[Hint : Let $t < \alpha$ and set $A_t = \bigcup\{I_t : t < \alpha\} \cup \bigcup\{J_s : s < t\}$. Show that $I_t + J_t$ is indiscernible over A_t . Also, as J_t realizes some I_t -invariant type over A_t , any sequence in A_t indiscernible over I_t remains indiscernible over $I_t + J_t$. The result follows by induction on $t < \alpha$.]

4.2 Dp-ranks

Definition 4.11. Let p be a partial type over a set A , and let κ be a (finite or infinite) cardinal. We say $\text{dp-rk}(p, A) < \kappa$ if for every family $(I_t : t < \kappa)$ of mutually indiscernible sequences over A and $b \models p$, there is $t < \kappa$ such that I_t is indiscernible over Ab .

If $a \in \mathcal{U}$, then $\text{dp-rk}(a/A)$ stands for $\text{dp-rk}(\text{tp}(a/A), A)$.

We implicitly allow the sequences I_t to be sequences of infinite tuples. However, if we impose that they be sequences of finite tuples, then we obtain an equivalent definition. This is because if an indiscernible sequence $I = (\bar{a}_i : i \in \mathcal{I})$ is not indiscernible over some b , then there are finite $a'_i \subseteq \bar{a}_i$ such that the sequence $(a'_i : i \in \mathcal{I})$ is indiscernible, but not indiscernible over b .

Remark 4.12. For p a type of a finite tuple, by Proposition 4.7, we always have $\text{dp-rk}(p, A) < |T|^+$.

We say that $\text{dp-rk}(a/A) = \kappa$ if $\text{dp-rk}(a/A) < \kappa^+$, but not $\text{dp-rk}(a/A) < \kappa$. Note that it may happen that we have $\text{dp-rk}(a/A) = \kappa$ for no value of κ . Consider for example the following situation: for every integer n , we can find a family $(I_t : t < n)$ of mutually indiscernible sequences over A , none of which is indiscernible over Aa , but we cannot find such a family of size \aleph_0 . We would then have $\text{dp-rk}(p, A) < \aleph_0$, but $\text{dp-rk}(p, A) \geq n$ for all n . One could probably write “ $\text{dp-rk}(a/A) = \aleph_0^-$ ” (as done for example in [2]), but we will not use that notation.

Lemma 4.13. Let p be a partial type over A and κ any cardinal. Let also $A \subseteq B$. Then $\text{dp-rk}(p, A) < \kappa \iff \text{dp-rk}(p, B) < \kappa$.

Proof. Assume that $\text{dp-rk}(p, A) < \kappa$. Let $(I_t : t < \kappa)$ be mutually indiscernible over B sequences and let $b \models p$. For $t < \kappa$, write $I_t = (a_i^t : i \in \mathcal{I}_t)$ and define the sequence $J_t = (a_i^t \wedge B : i \in \mathcal{I}_t)$. Then clearly, the sequences $(J_t : t < \kappa)$ are mutually indiscernible over A . By hypothesis, there is $t < \kappa$ such that J_t is indiscernible over Ab . This implies that I_t is indiscernible over Bb .

Conversely, assume that we have a witness of $\text{dp-rk}(p, A) \geq \kappa$. Namely, we have some $b \models p$ and a family $(I_t : t < \kappa)$ of sequences, mutually indiscernible over A such that no I_t is indiscernible over Ab . Composing by an automorphism, we may assume that the sequences are mutually indiscernible over B . Then we obtain a witness of $\text{dp-rk}(p, B) \geq \kappa$. \square

Remark 4.14. If $p \subseteq q$ are partial types over B , then clearly $\text{dp-rk}(p, B) < \kappa$ implies $\text{dp-rk}(q, B) < \kappa$. It follows that if $A \subseteq B$ and $a \in \mathcal{U}$, then $\text{dp-rk}(a/A) < \kappa$ implies $\text{dp-rk}(a/B) < \kappa$.

Proposition 4.15. *Let p be a partial type over A and let κ be any cardinal. Then we have $\text{dp-rk}(p, A) < \kappa$ if and only if for any family $(I_t : t \in X)$ of sequences, mutually indiscernible over A and any $b \models p$, there is $X_0 \subseteq X$ of size $< \kappa$ such that $\{I_t : t \in X \setminus X_0\}$ are mutually indiscernible over Ab .*

Proof. It is obvious that the property considered implies $\text{dp-rk}(p, A) < \kappa$. We show the converse.

Case 1: κ is infinite.

Assume that $\text{dp-rk}(p, A) < \kappa$ and assume that $(I_t : t \in X)$ and b give a counter-example to what we have to prove. Without loss, $X = \kappa$. We can build an increasing sequence $(\delta_t : t < \kappa)$ of ordinals, and a sequence $(\Delta_t : t < \kappa)$ of finite subsets of κ such that:

- for all $t < \kappa$, the sequence I_{δ_t} is not indiscernible over $\{b\} \cup \bigcup \{I_s : s \in \Delta_t\}$;
- for all $t < t' < \kappa$, $(\Delta_t \cup \{\delta_t\}) \cap (\Delta_{t'} \cup \{\delta_{t'}\}) = \emptyset$.

Let $B = A \cup \bigcup \{I_s : s \in \Delta_t, t < \kappa\}$. Then the sequences $\{I_{\delta_t} : t < \kappa\}$ are mutually indiscernible over B , and for each $t < \kappa$, I_t is not indiscernible over Bb . This contradicts $\text{dp-rk}(p, B) = \text{dp-rk}(p, A) < \kappa$.

Case 2: $\kappa = n + 1$ is finite.

Assume that $\text{dp-rk}(p, A) < n + 1$ and let $(I_t : t \in X)$ be sequences, mutually indiscernible over A . We may assume that X is finite and we show the result by induction on $|X|$.

If $|X| \leq n$, the result is obvious as we may take $X_0 = X$.

Assume that $|X| = n + k + 1$, and we have shown the result for sets of cardinality $\leq n + k$. Let $(I_t : t < n + k + 1)$ be mutually indiscernible over A and let $b \models p$. We may assume that the sequences I_t are endless. For $t < n + k + 1$, let p_t denote the limit type $\lim(I_t)$. Construct sequences $I'_t = (c_k^t : k < \omega)$, such that $c_k^t \models p_t \upharpoonright Ab \cup \{I_t : t \leq n + k\} \cup \{I'_s : s < t\} \cup \{c_l^t : l > k\}$ (so those are Morley sequences of each p_t , but read backwards).

Then the sequences $\{I_t + I'_t : t < n + k + 1\}$ are mutually indiscernible over A (see Exercise 4.10).

Let $B = A \cup \bigcup \{I'_t : t < n + k + 1\}$.

As $\text{dp-rk}(p, B) = \text{dp-rk}(p, A) \leq n$, there is some $t < n + k + 1$ such that I_t is indiscernible over Bb . Without loss, assume that $t = 0$.

By induction hypothesis, working over the base set AI_0 , there is some $X_0 \subset \{1, \dots, n + k\}$ of size at most n such that the sequences $\{I_t + I'_t : t \in \{1, \dots, n + k\} \setminus X_0\}$ are mutually indiscernible over AI_0b . Again, without loss, assume that the sequences $\{I_t + I'_t : 0 < t < k + 1\}$ are mutually indiscernible over AI_0b . If I_0 is indiscernible over $A \cup \bigcup \{I_t : 0 < t < k + 1\}$ then the sequences $\{I_t : t < k + 1\}$ are mutually indiscernible over Ab and we are done.

Otherwise, there is some formula $\phi(x, \bar{d}) \in L(Ab \cup \bigcup \{I_t : 0 < t < k + 1\})$ which witnesses that I_0 is not indiscernible over $Ab \cup \bigcup \{I_t : 0 < t < k + 1\}$. As the sequences $\{I_t + I'_t : 0 < t < k + 1\}$ are mutually indiscernible over AI_0b , we can move the parameters in ϕ belonging to some I_k to parameters in I'_k . Formally, there is some $\bar{d}' \equiv_{AI_0b} \bar{d}$ such that $\bar{d}' \in AI_0 \cup \bigcup \{I'_t : 0 < t' < k + 1\}$. Then $\phi(x, \bar{d}')$ witnesses the fact that I_0 is not indiscernible over Bb . This contradicts the assumption on I_0 . \square

We recall that the notation $\text{compl}(\mathcal{I})$ is used to denote the completion of the linear order \mathcal{I} . Recall also the definition of the equivalence relation $\sim_{\bar{c}}$ as defined before Proposition 3.16.

Theorem 4.16. *Let p be a partial type over A , and κ a cardinal (possibly finite). The following are equivalent:*

- (i) $\text{dp-rk}(p, A) < \kappa$;
- (ii)₀ *If $(I_t : t \in X)$ are mutually indiscernible over A and $b \models p$, then there is some $X_0 \subseteq X$ of size $< \kappa$ such that for $t \in X \setminus X_0$, all the members of I_t have the same type over Ab ;*
- (ii)₁ *Same as above, but change the conclusion to: for $t \in X \setminus X_0$, the sequence I_t is indiscernible over Ab ;*
- (ii)₂ *Same as above, but change the conclusion to: the sequences $\{I_t : t \in X \setminus X_0\}$ are mutually indiscernible over Ab ;*
- (iii)₀ *If $I = (a_i : i \in \mathcal{I})$ is a sequence (of possibly infinite tuples), there is some $\bar{c} \in \text{compl}(\mathcal{I})$, $|\bar{c}| < \kappa$ such that $i \sim_{\bar{c}} j \Rightarrow \text{tp}(a_i/Ab) = \text{tp}(a_j/Ab)$;*
- (iii)₁ *Same as above, with the conclusion that each $\sim_{\bar{c}}$ -class is indiscernible over Ab ;*

(iii)₂ Same as above, with the conclusion that the $\sim_{\bar{c}}$ -classes which are infinite are mutually indiscernible over Ab .

Proof. (i) \Rightarrow (ii)₂: This is exactly Proposition 4.15.

(ii)₂ \Rightarrow (ii)₁ \Rightarrow (ii)₀: Clear.

(ii)₀ \Rightarrow (i): Assume (ii)₀ and assume for a contradiction that there is some family $(I_t : t < \kappa)$ of sequences, mutually indiscernible over A and $b \models p$ such that for each t , I_t is not indiscernible over Ab . Without loss, each I_t is indexed by \mathbb{Q} : $I_t = (a_i^t : i \in \mathbb{Q})$. Then for each t , there are $\alpha_t < \beta_t \in \mathbb{Q}$ such that $\text{tp}(a_{\alpha_t}/AC_t a) \neq \text{tp}(a_{\beta_t}/AC_t a)$ where $C_t = \{a_i^t : i < \alpha_t \text{ or } i > \beta_t\}$. Set $I'_t = (a_i^t \wedge C_t : \alpha_t \leq i \leq \beta_t)$. Then the family $(I'_t : t < \kappa)$ contradicts (ii)₀.

(iii)₂ \Rightarrow (iii)₁ \Rightarrow (iii)₀: Clear.

(ii)₂ \Rightarrow (iii)₂: Assume (ii)₂ and let $I = (a_i : i \in \mathcal{I})$ be indiscernible over A , and $b \models p$. Without loss, \mathcal{I} is a complete order, so $\text{compl}(\mathcal{I}) = \mathcal{I}$. We may extend I to some sequence $J = (a_i : i \in \mathcal{J})$ where $\mathcal{I} \subset \mathcal{J}$ for \mathcal{J} some sufficiently large complete order and J is indiscernible over A . As \mathcal{I} is complete, to every point $x \in \mathcal{J}$ which lies in the convex hull of \mathcal{I} , there is a unique point $\pi(x) \in \mathcal{I}$ such that

$$(\forall i \in \mathcal{I})(i > \pi(x) \implies i \geq j) \wedge (i < \pi(x) \implies i \leq j).$$

As \mathcal{J} is a complete order, for every formula $\phi(x_1, \dots, x_n; y)$, with $|x_k| = |a_i|$ and $|y| = |b|$, there is some finite $\bar{d}_\phi \in \mathcal{J}^m$ such that for any $\bar{i}, \bar{j} \in \mathcal{J}^n$, we have $\bar{i} \sim_{\bar{d}_\phi} \bar{j} \implies \phi(a_{\bar{i}}; b) \leftrightarrow \phi(a_{\bar{j}}; b)$. Let \mathbf{d} be the union of (the ranges of) all tuples \bar{d}_ϕ for ϕ ranging over all formulas as above.

Increasing \mathcal{J} if necessary, we can find open disjoint intervals $(\mathcal{J}_i : i \in \mathcal{I} \cup \mathbf{d})$ such that $i \in \mathcal{J}_i$ for all $i \in \mathcal{I}$, and $d \in \mathcal{J}_d$ for all $d \in \mathbf{d}$.

Let E be the (convex) equivalence relation on \mathcal{J} defined by: jEj' if and only if, for all $i \in \mathbf{d}$, either $(j \in \mathcal{J}_i \text{ and } j' \in \mathcal{J}_i)$ or $(j < \mathcal{J}_i \text{ and } j' < \mathcal{J}_i)$ or $(j > \mathcal{J}_i \text{ and } j' > \mathcal{J}_i)$. Let \mathfrak{K} be the family of all E -classes. Then the sequences $\{(c_i : i \in X) : X \in \mathfrak{K}\}$ are mutually indiscernible over Ab . Furthermore, by construction of E , for each $X \in \mathfrak{K}$ which intersects the convex hull of \mathcal{I} , there is some $c_X \in \mathcal{I}$ such that $c_X = \pi(x)$ for all $x \in X$.

By (ii)₂, there is some $\mathfrak{K}_0 \subseteq \mathfrak{K}$ of size $< \kappa$ such that the sequences in $\{(c_i : i \in X) : X \in \mathfrak{K} \setminus \mathfrak{K}_0\}$ are mutually indiscernible over Ab . Let $\mathfrak{K}^* \subseteq \mathfrak{K}_0$ denote the set of sequences which intersect the convex hull of \mathcal{I} . Finally, let $\bar{c} = (c_X : X \in \mathfrak{K}^*)$. Then $|\bar{c}| < \kappa$.

We show that $\sim_{\bar{c}}$ has the required property. Let $\phi(x_1, \dots, x_n; y) \in L(A)$ be a formula. Let $i_1 < \dots < i_n \in \mathcal{I} \setminus \bar{c}$ and let $k \in \{1, \dots, n\}$. It is enough to

show that $\models \phi(a_{i_1}, \dots, a_{i_n}; b) \leftrightarrow \phi(a_{i_1}, \dots, a_{i'_k}, \dots, a_{i_n}; b)$ for any $i'_k \in \mathcal{I}$ such that $i_{k-1} < i'_k < i_{k+1}$ and $i'_k \sim_{\bar{c}} i_k$.

Assume this does not hold, as witnessed by some $i'_k > i_k$. Let

$$i_* = \inf\{j \in \mathcal{J} : j > i_k \text{ \& \& } \models \neg\phi(a_{i_1}, \dots, a_{i_n}; b) \leftrightarrow \phi(a_{i_1}, \dots, a_j, \dots, a_{i_n}; b)\}.$$

Then i_* belongs to the tuple \bar{d}_ϕ . In particular, i_* lands in the interior of some member of \mathfrak{K} . As $i_k \leq i_* \leq i'_k$ and $i_k \sim_{\bar{c}} i'_k$, i_* is in some sequence X of $\mathfrak{K} \setminus \mathfrak{K}_0$. Therefore the sequence $(a_i : i \in X)$ is indiscernible over $Ab \cup \{a_{i_k} : k \leq n, a_{i_k} \notin X\}$. This contradicts the definition of i_* .

(iii)₀ \Rightarrow (ii)₀: Let $(I_t : t < \alpha)$ be mutually indiscernible over A , and let $b \models p$. Write $I_t = (a_i^t : i \in \mathcal{I}_t)$, where the indexing orders \mathcal{I}_t are disjoint. Let \mathcal{I} denote the linear order $\bigoplus_{t < \alpha} \mathcal{I}_t$. We may expand the family $(I_t : t < \alpha)$ into a family $(J_t : t < \alpha)$ of sequences, mutually indiscernible over A such that $J_t = (a_i^t : i \in \mathcal{I})$ (in particular, I_t is a subsequence of J_t). Let $J = (a_i^{0 \wedge} a_i^{1 \wedge} \dots : i \in \mathcal{I})$. Then J is an indiscernible sequence. By assumption (iii)₀, there is some $\bar{c} \subseteq \text{compl}(I)$ of size $< \kappa$ such that

$$i \sim_{\bar{c}} j \implies \text{tp}(a_i^{0 \wedge} a_i^{1 \wedge} \dots / Ab) = \text{tp}(a_j^{0 \wedge} a_j^{1 \wedge} \dots / Ab).$$

Let $X_0 \subset \alpha$ be the set of $t < \alpha$ such that $X_0 \cap \mathcal{I}_t \neq \emptyset$. Then $|X_0| < \kappa$ and X_0 satisfies the requirement of (ii)₀. \square

EXAMPLE 4.17. In (iii)₂, we insist that the *infinite* classes are mutually indiscernible. If the indiscernible sequence I is densely ordered, then all classes are either infinite or reduced to a point. So the conclusion says nothing about the latter ones. We could wonder whether we can ask for the infinite classes to be mutually indiscernible over $(A$ and) the finite ones. This turns out to be an even stronger property, which we could call $\text{dp-rk}^+(p, A) < \kappa$ (because it is linked with what Shelah calls *strongly*⁺-dependent). We will not go into this, but only give an example where the two notions differ.

Let $T = \text{ACVF}$, the theory of algebraically closed valued fields (see Section A.2). Let $(a_i : i \in \mathbb{Q})$ be an indiscernible sequence of elements such that $v(a_i) < v(a_j)$ for all $i < j$. Then we can find a point b such that $v(b - (a_0 + \dots + a_n)) > v(a_n)$ for all $n < \omega$ (we then necessarily have $v(b - (a_0 + \dots + a_n)) = v(a_{n+1})$). In other words, b is a pseudo-limit of the sequence $(a_0 + \dots + a_n : n < \omega)$.

Then the point b splits the sequence I into three pieces: $(a_i : i < 0)$, (a_0) and $(a_i : i > 0)$. The two infinite ones are mutually indiscernible over b . However, the sequence $(a_i : i > 0)$ is not indiscernible over $a_0 b$. In fact that

tuple breaks it again into three pieces: $(a_i : 0 < i < 1)$, (a_1) and $(a_i : i > 1)$. And we can go on.

(In Shelah's terminology, this shows that ACVF is not strongly⁺-dependent. However it is strongly dependent, even dp-minimal, as defined below.)

Proposition 4.18. *Let $a, b \in \mathcal{U}$ and A a small set of parameters and κ_1, κ_2 two cardinals such that $\text{dp-rk}(a/Ab) < \kappa_1$ and $\text{dp-rk}(b/A) < \kappa_2$, then $\text{dp-rk}(a, b/A) < \kappa_1 + \kappa_2 - 1$.*

Proof. Here $\kappa - 1$ is equal to κ when κ is infinite.

We use condition (ii)₂. Let $(I_t : t \in X)$ be mutually indiscernible over A . There is some $X_0 \subset X$ of size $< \kappa_1$ such that the sequences $I_t, t \in X \setminus X_0$ are indiscernible over Ab . Then there is some $X_1 \subseteq X \setminus X_0$ of size $< \kappa_2$ such that the sequences $I_t, t \in X \setminus (X_0 \cup X_1)$ are indiscernible over Aab . \square

Note in particular that the hypothesis of the proposition are satisfied if $\text{dp-rk}(a/A) < \kappa_1$ and $\text{dp-rk}(b/A) < \kappa_2$.

Another characterization of dp-ranks is by Shelah-style arrays.

Definition 4.19. Let $p(y)$ be a partial type over a set A . We define $\kappa_{\text{ict}}(p, A)$ as the minimal κ such that the following does not exist:

- formulas $\phi_\alpha(x_\alpha; y)$;
- an array $(a_i^\alpha : i < \omega, \alpha < \kappa)$ of tuples, with $|a_i^\alpha| = |x_\alpha|$;
- for every $\eta : \kappa \rightarrow \omega$, a tuple b_η such that we have

$$\models \phi(a_i^\alpha; b_\eta) \iff \eta(\alpha) = i.$$

(Such a family of tuples and formulas will be called an *ict-pattern* for p . The cardinal κ is the *length* of the pattern).

Proposition 4.20. *For any partial type p over A and cardinal κ , we have $\text{dp-rk}(p, A) < \kappa$ if and only if $\kappa_{\text{ict}}(p, A) \leq \kappa$.*

Proof. Assume there is an ict-pattern for p of length κ . Then by expanding and extracting, we may assume that the rows $(a_i^\alpha : i < \omega)$, $\alpha < \kappa$, are mutually indiscernible over A . Then for any η , b_η witnesses that $\text{dp-rk}(p, A) \geq \kappa$.

Conversely, assume that $\text{dp-rk}(p, A) \geq \kappa$. By condition (ii)₀ of Theorem 4.16 we can find:

- sequences $(I_t : t < \kappa)$, mutually indiscernible over A , without loss $I_t = (a_i^t : i < \omega)$ are sequences of finite tuples;

- formulas $\phi_t(x_t; y)$;
- a tuple $b \models p$ such that $\models \phi_t(a_0^t; b) \wedge \neg \phi_t(a_1^t; b)$.

Let $\psi_t(x_t, x'_t; y) = \phi_t(x_t; y) \wedge \neg \phi_t(x'_t; y)$. For $t < \kappa$, let $J_t = (a_{2i}^t \hat{a}_{2i+1}^t : i < \omega)$. By NIP, at most finitely many tuples $a \hat{a}'$ from J_t satisfy $\psi_t(x_t, x'_t; b)$. Removing some if necessary, we may assume that $a_0^t \hat{a}_1^t$ is the only element of J_t for which that formula holds. By mutual indiscernability of the sequences $\{J_t : t < \kappa\}$, for every $\eta : \kappa \rightarrow \omega$, we can find some b_η such that, for all $t < \kappa$,

$$\models \psi_t(a_{2i}^t, a_{2i+1}^t; b_\eta) \iff \eta(t) = i.$$

This shows that $\kappa_{ict}(p, A) > \kappa$. \square

Note that the difference between dp-rk and κ_{ict} is only due to the convention chosen for the definition of dp-rank. One argument in favor of this definition, as opposed to that of κ_{ict} , is that 1-types in dp-minimal theories are of rank 1 and not 2.

4.3 Strongly dependent theories

Definition 4.21. The NIP theory T is *strongly dependent* if for any finite tuple of variables x , we have $\text{dp-rk}(x = x, \emptyset) < \aleph_0$.

EXAMPLE 4.22. Let $\mathcal{L} = \{E_i : i < \omega\}$ where the E_i 's are binary predicates. Let T be the \mathcal{L} -theory stating that each E_i defines an equivalence relation with infinitely many classes, each of which is infinite. Consider $T_1 \supset T$ stating that $(\forall xy)xE_{i+1}y \rightarrow xE_iy$, and each E_i -class is split into infinitely many E_{i+1} classes.

Consider also $T_2 \supset T$ stating that the E_i 's are cross-cutting, that is: given a_0, \dots, a_{n-1} , there is a such that $aE_k a_k$ for every $k < n$.

Then T_1 is strongly dependent, but not T_2 (and note that both are stable).

Remark 4.23. If T is a superstable theory, then it is strongly dependent, but the converse does not hold as witnessed by T_1 in the previous example.

Proposition 4.24. *If for every variable x with $|x| = 1$, we have $\text{dp-rk}(x = x, \emptyset) < \aleph_0$, then T is strongly dependent.*

Proof. This follows immediately from Proposition 4.18. \square

An extreme case of strongly dependent theories are dp-minimal theories.

Definition 4.25. The theory T is *dp-minimal*, if $\text{dp-rk}(x = x, \emptyset) = 1$, for x a singleton.

EXAMPLE 4.26. The following theories are dp-minimal (see Appendix A or [28]):

- any o-minimal theory;
- any C -minimal theory (for example, ACVF, the theory of algebraically closed valued fields;
- the p -adics;
- the theory of $(\mathbb{Z}, 0, 1, +, \leq)$.

We point out some link with honest definitions. First an easy lemma.

Lemma 4.27. *Let $A \subset \mathcal{U}$. Let $(p_i : i < \omega)$ be a family of A -invariant types and let $(c_i, d_i : i < \omega)$ and $b \in \mathcal{U}$ be finite tuples such that:*

- $c_i \models p_i \upharpoonright Ac_{<i}d_{<i}$;
- $d_i \models p_i \upharpoonright Ac_{\leq i}d_{<i}$;
- $\text{tp}(c_i/Ab) \neq \text{tp}(d_i/Ab)$.

Then T is not strongly dependent.

Proof. Build sequences $I_i = (e_j^i : j < \omega)$ for $i < \omega$ such that I_i is a Morley sequence of p_i over $AI_{<i}$. Then we see that $\text{tp}((c_i, d_i : i < \omega)/A) = \text{tp}((e_0^i, e_1^i : i < \omega)/A)$. Therefore, composing by an automorphism, we may assume that $e_0^i = c_i$ and $e_1^i = d_i$ for all $i < \omega$. Then the sequences $\{I_i : i < \omega\}$ are mutually indiscernible over A , but none of them remains indiscernible over Ab . This implies that T is not strongly dependent. \square

Proposition 4.28. *Let T be strongly dependent. Let $M \models T$, $A \subseteq M$, $b \in M$ a finite tuple, and $(M, A) \prec (M', A')$ some $|M|^+$ -saturated extension. Then there is $A_0 \subseteq A'$ finite such that every formula $\phi(x; b) \in L(b)$ has an honest definition over A with parameters in AA_0 .*

The conclusion means that we can find $\psi(x; c) \in L(AA_0)$ such that $\phi(A; b) \subseteq \psi(A'; c) \subseteq \phi(A'; b)$.

Proof. For x a finite tuple of variables, let $S_x^A \subset S_x(\mathcal{U})$ be the set of global types, finitely satisfiable in A . By Lemma 4.27, we can build a maximal sequence $(c_i, d_i : i < N)$ such that there are types $p_i \in S_{x_i}^A$ with:

- $c_i \models p_i \upharpoonright Ac_{<i}d_{<i}$;
- $d_i \models p_i \upharpoonright Ac_{\leq i}d_{<i}$;

– $\text{tp}(c_i/Ab) \neq \text{tp}(d_i/Ab)$.

Set $A_0 = \{c_i, d_i : i < N\}$.

Let now $\phi(x; b)$ be given. Then for any type $q \in S_x^A$, there is a truth value $\epsilon_q \in \{0, 1\}$ such that for any realization a of $q|_{AA_0}$ in A' , we have $\models \phi(a; b)^{\epsilon_q}$. By saturation of the pair (M', A') , there is a formula $\psi_q(x) \in q|_{AA_0}$ such that $(M', A') \models (\forall x \in \mathbf{P}) \psi_q(x) \rightarrow \phi(x; b)^{\epsilon_q}$.

We then conclude exactly as in the proof of Theorem 3.1. (Namely, we extract from $\{\psi_q(x) : \epsilon_q = 1\}$ a finite subcover, and take the union of those formulas.) \square

Note that the converse does not hold, since any stable theory satisfies the conclusion of the proposition, but not all stable theories are strongly dependent.

We end this section with a example of an application to definable groups.

We say that a group G is of finite exponent if there exists $n < \omega$ such that $g^n = e$ for all $g \in G$.

Proposition 4.29. *Let G be a dp-minimal group. Then there is a definable, abelian, normal subgroup H such that the quotient G/H is of finite exponent.*

Proof. Claim: For any two definable subgroups H and K of G , one of $[H : H \cap K]$ or $[K : H \cap K]$ is finite.

Proof: Assume not. Then we can find two sequences $(a_i : i < \omega)$ of points of K and $(b_i : i < \omega)$ of points of H such that the cosets $\{a_i H : i < \omega\}$ are distinct, as well as the cosets $\{K b_i : i < \omega\}$. Note that for any $i, j < \omega$, there is a point in $a_i H \cap K b_j$, namely $a_i b_j$. And as two distinct cosets of the same group are disjoint, we have $a_i b_j \in a_{i'} H \cap K b_{j'} \iff (i, j) = (i', j')$. By Proposition 4.20, this shows that $\text{dp-rk}(a_i b_j, \emptyset) \geq 2$, contradicting dp-minimality.

It follows that if H and K are two definable subgroups, $a \in H$ and $b \in K$, then there is n such that either $a^n \in K$ or $b^n \in H$. If $a \in G$, we let $C(a) = \{g \in G : ga = ag\}$. Then applying this to the uniform family $\{C(a) : a \in G\}$ of definable groups yields: there is some integer n such that for any two elements $a, b \in G$, there is $k \leq n$ such that either $a^k \in C(b)$ or $b^k \in C(a)$. In particular, if $N = n!$, then a^N and b^N commute.

Let $H = C(C(G^N))$, then H has the required properties. \square

Historical remarks

Strong dependence was defined by Shelah in [83] as a tentative analog of superstable, and further studied in [77]. In that latter paper, he defines a whole

family of “dp-ranks” which depend not only on a type p and a set I , but also of a set B of indiscernible sequences, or pairs of points, “split” by p . This allows to make constructions where the rank drops. He uses them to prove that in a strongly dependent theory, any long enough sequence contains an infinite indiscernible subsequence. The notion of dp-rank we present here is a simplification of those ranks, it is also similar to Shelah’s κ_{ict} , but restricted to a type. It was first explicitly defined by Usvyatsov in [89] and it coincides in NIP theories with what Adler had previously called *burden* in [2].

Proposition 4.15 and the additivity of dp-rank that follows from it is from [49]. (In which the convention for $\text{dp-rk}(p)$ is slightly different. To be precise, $\text{dp-rk}(p) < \kappa$ for us is equivalent to $\text{dp-rk}(p) \leq \kappa - 1$ in the sense of [49]. In particular, for infinite cardinals, that latter definition agrees with κ_{ict} , but not for finite ones.) The equivalent characterizations given in Theorem 4.16 are from [83], except for (ii)₂ and (iii)₂ which make use of Proposition 4.15.

The notion of a dp-minimal theory originates in Shelah’s work, but was precisely defined and studied by Onshuus and Usvyatsov in [64].

Proposition 4.28 is new. Proposition 4.29 is from [88].

Chapter 5

Forking

Given two sets $A \subseteq B$ and a type $p \in S(A)$, we want to define a notion of a free extension of p to a type over B . When $A = M$ is a model, the work done in Chapter 2 provides us with a natural option: an extension of p is free if it extends to a global M -invariant type. However, it is not immediately clear how to generalize this to arbitrary base sets.

The idea of forking is to give such a definition (or rather a definition of the opposite: the free extensions are the non-forking ones). This definition makes sense in any theory, and satisfies different kinds of properties depending on the characteristics of the theory. For example, in *simple* theories, non-forking is an *independence relation* which means in particular that it satisfies symmetry and transitivity (see for example [18]). In NIP theories, this is not true, but other properties hold. In particular a type has only boundedly many non-forking extensions, and over models, non-forking coincides with invariance. As we will see, over an arbitrary set A , it is equivalent to $\text{bdd}(A)$ -invariance.

Even though the general definition of forking coincides, for NIP theories, with an apparently simple one, experience shows that it is useful to come back to it from time to time.

5.1 Bounded equivalence relations

In the first three sections of this chapter, we do not assume NIP.

An equivalence relation E between tuples of \mathcal{U} is *A-invariant* if whenever $a \equiv_A a'$ and $b \equiv_A b'$, then $aEb \iff a'Eb'$. The relation xEy is type-definable over A if it is defined by a partial type $\pi(x; y)$ over A .

Proposition 5.1. (*T any theory*)

Let $A \subset \mathcal{U}$ and let E be an A -invariant equivalence relation on \mathcal{U} . Then the following are equivalent:

- (i) the set \mathcal{U}/E is bounded (i.e., of size $< \bar{\kappa}$);
- (ii) $|\mathcal{U}/E| \leq 2^{|A|+|T|}$;
- (iii) for any A -indiscernible sequence $(a_i : i < \omega)$ and $i, j < \omega$, we have $a_i E a_j$;
- (iv) for any model $M \supseteq A$ and $a \equiv_M b$, we have $a E b$.

Proof. (ii) \Rightarrow (i): Clear.

(iv) \Rightarrow (ii): Clear.

(iii) \Rightarrow (iv): Let $a \equiv_M b$ and take p to be a coheir of $\text{tp}(a/M)$. In particular, p is M -invariant. Let $(a_i : i < \omega)$ be a Morley sequence of p over Mab . Then both $(a) + (a_i : i < \omega)$ and $(b) + (a_i : i < \omega)$ are indiscernible sequences. Therefore by (iii), we have $a E a_0 E b$.

(i) \Rightarrow (iii): Assume there is some A -indiscernible sequence $(a_i : i < \omega)$ such that $\neg(a_i E a_j)$ for $i \neq j$. Then we may increase the sequence to one of size $\bar{\kappa}$. This contradicts (i). \square

An A -invariant equivalence relation E is *bounded* if it satisfies one of the equivalent conditions above.

Definition 5.2. We say that two tuples a and b have the same Lascar strong-type over A if we have $a E b$ for any A -invariant bounded equivalence relation. We write $\text{Lstp}(a/A) = \text{Lstp}(b/A)$.

Lemma 5.3. *The relation “having same Lascar strong-type over A ” is the finest bounded A -invariant equivalence relation.*

It is the transitive closure of the relation $R_A(a; b)$ defined to hold if there is a model $M \supseteq A$ such that $a \equiv_M b$.

It is also the transitive closure of the relation $S_A(a; b)$ defined to hold if (a, b) is the beginning of some A -indiscernible sequence.

Proof. This is clear from the previous proposition. \square

We let $\text{Aut}f(\mathcal{U}/A)$ be the set of automorphisms of \mathcal{U} fixing Lascar strong-types over A . In other words, it is the set of $\sigma \in \text{Aut}(\mathcal{U}/A)$ such that $\text{Lstp}(a/A) = \text{Lstp}(\sigma(a)/A)$ for all a . It is also the subgroup of $\text{Aut}(\mathcal{U})$ generated by the groups $\text{Aut}(\mathcal{U}/M)$ for M a model containing A .

We say that a global type p is Lstp_A -invariant if it is invariant under $\text{Aut}(\mathcal{U}/A)$. This is equivalent to saying that if $(a_i : i < \omega)$ is an A -indiscernible sequence in \mathcal{U} and $a \models p$, then $(a_i : i < \omega)$ is indiscernible over Aa .

Definition 5.4. We say that two tuples a and b have the same compact strong-type over A (or Kim-Pillay strong-type), written $\text{KP-stp}(a/A) = \text{KP-stp}(b/A)$ if aEb holds for every bounded type-definable over A equivalence relation E .

Therefore “equality of compact strong-types over A ” is the finest type-definable over A bounded equivalence relation.

We let $\text{Aut}(\mathcal{U}/\text{bdd}(A))$ be the set of automorphisms of \mathcal{U} fixing all compact strong types over A . We say that a global type p is $\text{bdd}(A)$ -invariant if it is invariant under $\text{Aut}(\mathcal{U}/\text{bdd}(A))$.

Note that Lascar strong-types refines compact strong-types refines types. If $A = M$ is a model, then all those notions coincide.

Proposition 5.5. *Let p be a type over A and let $R(x; y)$ be an A -invariant relation whose restriction to $p(\mathcal{U})^2$ is an equivalence relation with boundedly many classes, then:*

- (i) *there is an A -invariant bounded equivalence relation E whose restriction to $p(\mathcal{U})^2$ coincides with R ;*
- (ii) *if furthermore R is type-definable, then there is a type-definable (over A) bounded equivalence relation E whose restriction to $p(\mathcal{U})^2$ refines R .*

Proof. Point (i) is easy: simply take $E(x; y)$ to be $(R(x; y) \wedge p(x) \wedge p(y)) \vee (\neg p(x) \wedge \neg p(y))$.

Assume now that R is type definable. Let $(a_i : i < \alpha)$ be a family of representatives of each class of $p(\mathcal{U})/R$. Let $E(x; y)$ be the relation stating that there are $(a'_i, a''_i : i < \alpha)$, all satisfying p such that $R(a'_i, a_i) \wedge R(a''_i, a_i)$ holds for each $i < \alpha$ and $\text{tp}(x(a'_i)_{i < \alpha}/A) = \text{tp}(y(a''_i)_{i < \alpha}/A)$. Then E is type-definable over $A(a_i)_{i < \alpha}$. However, choosing another set of representatives yields the same relation E , therefore E is A -invariant and thus type-definable over A . It is clear that E is an equivalence relation. If x and y have the same type over $A(a_i)_{i < \alpha}$, then they are E -equivalent (take $a'_i = a''_i = a_i$), therefore E is bounded. Finally, if x and y satisfy p and are E -equivalent, then this is witnessed by some $(a'_i, a''_i : i < \alpha)$. For some $i < \alpha$, we have $R(x; a'_i)$, therefore also $R(y; a''_i)$ holds and then $R(x; y)$ holds. Thus E has the required properties. \square

5.2 Forking

Definition 5.6. Let $A \subset \mathcal{U}$.

(i) A formula $\phi(x; b) \in L(\mathcal{U})$ *divides* over A if there is an A -indiscernible sequence $(b_i : i < \omega)$ where $b_0 = b$ such that the partial type $\{\phi(x; b_i) : i < \omega\}$ is inconsistent.

(ii) A partial type $\pi(x)$ *divides* over A if it implies (equivalently, contains) a formula which divides over A ;

(iii) A partial type $\pi(x)$ *forks* over A if it implies a finite disjunction $\bigvee_{i < n} \phi_i(x; b^i)$ of formulas, such that $\phi_i(x; b^i)$ divides over A for each $i < n$.

The point of considering forking instead of dividing lies in observation 4 below.

Fact 5.7. 1. If $p(x)$ is a consistent complete type over A , then it does not divide over A . (However, it may fork over A , see Example 5.8.)

2. If the partial type $\pi(x)$ forks over A , then some finite $\pi_0(x) \subseteq \pi(x)$ forks over A .

3. If $p(x)$ is a global complete type, then it does not fork over A if and only if it does not divide over A .

4. Let $A \subseteq B$ and $\pi(x)$ be a partial type over B , then $\pi(x)$ does not fork over A if and only if it has an extension to a global complete type which does not divide (equiv. fork) over A .

5. If $A \subseteq B \subseteq C$ and $a \in \mathcal{U}$, then if $\text{tp}(a/C)$ does not divide over A , it does not divide over B .

6. If $p(x)$ is a global Lstp_A -invariant type, then it does not divide (equiv. fork) over A .

Proof. 1. If $\pi(x)$ implies a formula $\phi(x; b_0)$ and $(b_i : i < \omega)$ is an A -indiscernible sequence, then $\pi(x)$ implies all formulas $\phi(x; b_i)$, $i < \omega$. As $\pi(x)$ is consistent, it follows that $\phi(x; b_0)$ does not divide over A .

2. Clear by compactness.

3. Assume that $p(x)$ implies a finite disjunction of formulas $\bigvee_{i < n} \phi_i(x; b)$, each of which divides over A . There is some finite $C_0 \subseteq \mathcal{U}$ such that $p'(x) = p(x)|_{C_0}$ already implies this disjunction. Let $C = C_0 \cup A$. Then we may find

a tuple $b' \in \mathcal{U}$, $b' \equiv_C b$. Then $p'(x)$ implies $\bigvee_{i < n} \phi_i(x; b')$, hence one of the formulas $\phi_i(x; b')$ is implied by $p(x)$, and $p(x)$ divides over A .

4. Let $\pi(x)$ be a partial type over A . Let $\Sigma(x)$ be the set of formulas $\neg\phi(x; b)$ where $\phi(x; b) \in L(\mathcal{U})$ divides over A . If $\pi(x) \cup \Sigma(x)$ is consistent, then it extends to a complete type over \mathcal{U} which does not divide over A . Otherwise, by compactness, there is some finite part $\Sigma_0(x)$ of $\Sigma(x)$ such that $\pi(x) \cup \Sigma_0(x)$ is consistent. Hence $\pi(x)$ forks over A .

5. Clear from the definition.

6. Assume that $p(x) \vdash \phi(x; b)$ for some $\phi(x; b) \in L(\mathcal{U})$, and let $(b_i : i < \omega)$ be an A -indiscernible sequence with $b_0 = b$. Then as $p(x)$ is $Lstp_A$ -invariant, $p(x) \vdash \phi(x; b_i)$ for each $i < \omega$. Hence as p is consistent, so is the conjunction $\bigwedge_{i < \omega} \phi(x; b_i)$. \square

EXAMPLE 5.8. We give an example of a type $p \in S(A)$ which forks over A .

Let \mathbb{U} be the usual unit circle in the plane and define a ternary relation $R(x, y, z)$ on \mathbb{U} which holds if and only if x and z are diametrically opposite, or y lies in the (closed) small arc between x and z . Consider the structure $M = (\mathbb{U}; R)$. Let $p \in S_1(\emptyset)$ be the unique 1-type over \emptyset . We claim that p forks over \emptyset .

To see this, let $b, c, d \in \mathbb{U}$ divide the circle into three small arcs. Then $p \vdash R(b, x, c) \vee R(c, x, d) \vee R(d, x, b)$. It is enough to show that say $R(b, x, c)$ divides over \emptyset . We can find an indiscernible sequence $(b_i, c_i : i < \omega)$ such that $(b_0, c_0) = (b, c)$ and the points $b_0, c_0, b_1, c_1, \dots$ lie in that order on the circle. Then the formula $R(b_0, x, c_0) \wedge R(b_1, x, c_1)$ is inconsistent. This finishes the proof.

Note that this theory is NIP, since it is interpretable in RCF.

Lemma 5.9. *Let $A, B \subset \mathcal{U}$ and π a partial type over B . Then the following are equivalent:*

- (i) π does not divide over A ;
- (ii) for every A -indiscernible sequence $I = (c_i : i < \omega)$ with $c_0 \in B$, there is a $\models \pi$ such that I is indiscernible over Aa ;

Proof. It is clear that (ii) implies (i).

Conversely, assume (i) and let $I = (c_i : i < \omega)$ with $c_0 \in B$ be an A -indiscernible sequence. Increasing I if necessary, we may assume that c_0 enumerates B . Write $\pi(x) = \pi(x; c_0)$. By the non-dividing assumption, the partial type $\{\pi(x; c_i) : i < \omega\}$ is consistent. Let a' realize it. By Ramsey, we may find an Aa' -indiscernible sequence $(c'_i : i < \omega)$ which is based on $(c_i : i < \omega)$ over

Aa' . Then $\text{tp}(c'_0/Aa') = \text{tp}(c_0/Aa')$. Let $f \in \text{Aut}(\mathcal{U}/B)$ send $(c'_i : i < \omega)$ to $(c_i : i < \omega)$ and set $a = f(a')$. Then the sequence I is indiscernible over Aa . \square

Lemma 5.10. (*T any theory*) Let $A \subseteq B \subset \mathcal{U}$ and $a, b \in \mathcal{U}$. Assume that $\text{tp}(a/B)$ does not fork over A and $\text{tp}(b/Ba)$ does not fork over Aa , then $\text{tp}(a, b/B)$ does not fork over A .

Proof. Let M be some $|B|^+$ -saturated model containing B . As the type $\text{tp}(a/B)$ does not fork over A , it has an extension p' over M which does not fork over A . Let a' realize p' . There is an automorphism σ fixing B and sending a to a' . Replacing (a, b) by $(\sigma(a), \sigma(b))$ we may assume that $\text{tp}(a/M)$ does not fork over A . Similarly, as the type $\text{tp}(b/Ba)$ does not fork over Aa , it has an extension q' to a type over Ma which does not fork over Aa . Replacing b by a realization of q' , we may assume that $\text{tp}(b/Ma)$ does not fork over Aa .

Now it is enough to prove that $\text{tp}(a, b/M)$ does not divide over A . So assume that it does divide and take $\phi(x, y; c) \in \text{tp}(a, b/M)$ and an A -indiscernible sequence $(c_i : i < \omega)$, $c_0 = c$ such that $\{\phi(x, y; c_i) : i < \omega\}$ is inconsistent. By Lemma 5.9, there is an automorphism f fixing Ac pointwise such that $(c_i : i < \omega)$ is indiscernible over $f(a)$. Then $\text{tp}(f(b)/Af(a)c)$ does not divide (indeed does not fork) over $Af(a)$. Therefore there is an automorphism g fixing $Af(a)c$ such that $(f(a)c_i : i < \omega)$ is indiscernible over $g(f(b))$. We then have $\models \phi(f(a), g(f(b)); c_i)$ for all $i < \omega$. Contradiction. \square

Corollary 5.11. (*T any theory*) Assume that for every $A \subset \mathcal{U}$ and $p \in S_1(A)$, p does not fork over A , then for every A , no $p \in S_n(A)$ forks over A .

Proposition 5.12. (*T is NIP*) Let $A \subset \mathcal{U}$, and let $p \in S(\mathcal{U})$ be a global type. Then p does not fork over A if and only if it is Lstp_A -invariant.

Proof. Right to left has already been observed and holds in any theory.

We show left to right. Let $p \in S(\mathcal{U})$ and assume that p is not Lstp_A -invariant. This implies that there is some A -indiscernible sequence $(c_i : i < \omega)$ and some formula $\phi(x; y)$ such that $p \vdash \phi(x; c_0) \wedge \neg\phi(x; c_1)$. Consider the A -indiscernible sequence of pairs $(c_{2i}c_{2i+1} : i < \omega)$. The partial type $\{\phi(x; c_{2i} \wedge \neg\phi(x; c_{2i+1}))\}$ is inconsistent by NIP. This implies that the formula $\phi(x; c_0) \wedge \neg\phi(x; c_1)$ divides over A . \square

Corollary 5.13. (*T is NIP*) Let $M \prec \mathcal{U}$, and p be a global type. Then p does not fork over M if and only if p is M -invariant.

Let $p \in S(\mathcal{U})$ be a global type. If p is M -invariant, for some small M , then it does not fork over M . Conversely, if p does not fork over some small $A \subset \mathcal{U}$, then it is M -invariant for any model M containing A . So for a global type p the properties “ p is invariant over some small $M \subset \mathcal{U}$ ” and “ p does not fork over some small $A \subset \mathcal{U}$ ” are equivalent. We call such types *global invariant types*.

Corollary 5.14. (*T is NIP*) *Let $p \in S(A)$ be a type in a finite number of variables. Then p has at most $2^{|A|+|T|}$ non-forking global extensions.*

Proof. Take M a model containing A such that $|M| = |A| + |T|$. If q is a non-forking extension of p , then q does not fork over M , therefore it is M -invariant. We have seen in the proof of Proposition 2.31 that there are only $2^{|M|}$ global M -invariant types. \square

5.3 $\text{bdd}(A)$ -invariance

In this section, we assume that T is NIP.

Proposition 5.15. *Let $p \in S(\mathcal{U})$ be non-forking over A . Then the equivalence relation $\text{Lstp}(x/A) = \text{Lstp}(y/A)$ restricted to realizations of $p|_A$ is type-definable.*

Indeed, if $c, d \models p|_A$, then: $\text{Lstp}(c/A) = \text{Lstp}(d/A)$ if and only if there is some sequence I such that both $(c) + I$ and $(d) + I$ realize $p^{(\omega)}|_A$.

Proof. It is enough to prove the second statement.

\Leftarrow : Clear

\Rightarrow : Let c, d realize $p|_A$, and assume that $\text{Lstp}(c/A) = \text{Lstp}(d/A)$. Conjugating by an automorphism over A , we may assume that $\text{Lstp}(c/A) = \text{Lstp}(a_0/A)$, for $a_0 \models p$. We let $I \models p^{(\omega)}|_{Acd a_0}$. Then $a_0 + I$ is a realization of $p^{(\omega)}|_A$. As $p^{(\omega)}$ does not fork over A , by NIP it is Lascar-invariant over A , therefore both $(c) + I$ and $(d) + I$ realize $p^{(\omega)}$. \square

Corollary 5.16. *If $p \in S(A)$ does not fork over A , and $a, b \models p$ satisfy $\text{KP-stp}(a/A) = \text{KP-stp}(b/A)$, then $\text{Lstp}(a/A) = \text{Lstp}(b/A)$.*

In particular, if no type forks over its base, then compact strong types coincide with strong types. (We say that T is G -compact).

Proof. Let $p \in S(A)$ be non-forking over A . By the previous proposition, the relation $R(c; d) \equiv \text{Lstp}(c/A) = \text{Lstp}(d/A)$ is type-definable on realizations of p . By Proposition 5.5, there is a type-definable bounded equivalence relation E which refines R on p (and must therefore be equal to R by definition of Lascar strong types). If $c, d \models p$ have the same KP-strong type, then they are E -equivalent and thus R -equivalent. \square

We now prove an analog of Proposition 2.27 for types non-forking over A .

Proposition 5.17. *Let $p, q \in S(\mathcal{U})$ be two global types, both non-forking over A and let $\bar{a} \models p^{(\omega)}$ and $\bar{b} \models q^{(\omega)}$. Assume that $\text{Lstp}(\bar{a}/A) = \text{Lstp}(\bar{b}/A)$, then $p = q$.*

Proof. The proof is essentially the same as that of Proposition 2.27, replacing types over A by Lascar strong types over A .

Assume that there is some $\phi(x; d) \in L(\mathcal{U})$ such that $p \vdash \phi(x; d)$ and $q \vdash \neg\phi(x; d)$. Pick any model $M \supseteq A \cup \{d\}$. We build inductively two sequence $(\bar{a}_n : n < \omega)$ and $(\bar{b}_n : n < \omega)$ such that for all n :

- $\bar{a}_n \models p^{(\omega)} \upharpoonright M\bar{a}_{<n}\bar{b}_{<n}$;
- $\bar{b}_n \models q^{(\omega)} \upharpoonright M\bar{a}_{\leq n}\bar{b}_{<n}$.

It is enough now to show that the concatenation $\bar{a}_0 + \bar{b}_0 + \bar{a}_1 + \bar{b}_1 + \dots$ is indiscernible, because then the formula $\phi(x; d)$ alternates infinitely often on that sequence, contradicting NIP.

We show by induction on n that $I_n := \bar{a}_0 + \bar{b}_0 + \dots + \bar{a}_n$ has the same Lascar strong type over A as a realization of $q^{(\omega(2n+1))}$. For $n = 0$, this is the hypothesis. Assume we know it for n . Then as $q^{(\omega)}$ is Lascar-invariant over A and I_n has the same Lascar strong type over A as $q^{(\omega(2n+1))}$, the sequence $I_n + \bar{b}_n$ realizes $q^{(\omega(2n+2))}|_A$. Therefore it is an indiscernible sequence. It follows by Lemma 5.3 that $I_n + \bar{b}_n$ has same Lascar strong type over A as a realization of $q^{(\omega(2n+2))}$, which is the same as that of a realization of $p^{(\omega(2n+2))}$. We then add \bar{a}_{n+1} to this sequence by the same argument and this finishes the induction. \square

Corollary 5.18. *Let $p \in S(\mathcal{U})$ be non-forking over A , then p is $\text{bdd}(A)$ -invariant.*

Proof. Let $\sigma \in \text{Aut}(\mathcal{U}/\text{bdd}(A))$, $\bar{a} \models p^{(\omega)}|_M$. Then

$$\text{KP-stp}(\bar{a}/A) = \text{KP-stp}(\sigma(\bar{a})/A).$$

As \bar{a} and $\sigma(\bar{a})$ are both realizations of $p^{(\omega)}|_A$, which is non-forking over A , Corollary 5.16 implies that $\text{Lstp}(\bar{a}/A) = \text{Lstp}(\sigma(\bar{a})/A)$. Therefore by Proposition 5.17, $p = \sigma(p)$. \square

5.4 NTP_2 and the broom lemma

Definition 5.19. (T any theory) We say that a formula $\phi(x; y)$ has TP_2 if there is an array $(b_i^t : i < \omega, t < \omega)$ of tuples of size $|y|$ and $k < \omega$ such that:

- for any $\eta : \omega \rightarrow \omega$, the conjunction $\bigwedge_{t < \omega} \phi(x; b_{\eta(t)}^t)$ is consistent;
- for any $t < \omega$, $\{\phi(x; b_i^t) : i < \omega\}$ is k -inconsistent.

We say that $\phi(x; y)$ is NTP_2 if it does not have TP_2 . We say that the theory T is NTP_2 if all formulas are NTP_2 .

Proposition 5.20. *If T is NIP, then it is NTP_2 .*

Proof. Assume that some formula $\phi(x; y)$ has TP_2 , as witnessed by some $k < \omega$ and array $(a_i^t : i < \omega, t < \omega)$. We show that $\text{VC}(\phi) = \infty$.

For simplicity, we assume that $k = 2$, and we explain afterwards how to handle the general case.

Let $A = \{a_0^t : t < \omega\}$ and $A_0 \subseteq A$ any subset. Let $\eta : \omega \rightarrow \omega$ such that

$$\eta(t) = \begin{cases} 0; & \text{if } t \in A, \\ 1 & \text{otherwise.} \end{cases}$$

By the properties of the array, there is some $b_{A_0} \models \bigwedge_{t < \omega} \phi(x; a_{\eta(t)}^t)$. Then using 2-inconsistency, we have

$$b_{A_0} \models \phi(x; a_0^t) \iff t \in A_0.$$

Thus the formula $\phi(x; y)$ has IP.

We indicate two ways to treat the case $k > 2$. One option is first to increase the array to $(a_i^t : i < \omega, t < \kappa)$, by compactness. Next using standard Ramsey argument we may assume that the sequences $\{(a_i^t : i < \omega) : t < \omega\}$ are mutually indiscernible. Then we conclude by Proposition 4.7. Another argument is by counting types. Let N be large enough and for any $s < \omega$ consider the finite set $A_s = \{a_i^t : i < N, t < s\}$. Then the properties of the array imply that there are at least $(N/k)^s$ ϕ -types over A_s . As $|A_s| = N \cdot s$, this contradicts Sauer's lemma 6.3. □

The class of NTP_2 theories contains both NIP theories and simple theories. Most of the content of this section and the next one goes through (with some slight modifications) for NTP_2 theories. However, we do not try to be optimal here and we will assume NIP throughout.

Notation 5.21. $a \perp_A b$ means that $\text{tp}(a/Ab)$ does not fork over A .

Note that this is not in general a symmetric relation.

Definition 5.22. An *extension base* is a set $A \subset \mathcal{U}$ such that no $p \in S(A)$ forks over A . Equivalently (using NIP), any $p \in S(A)$ has a global extension which is Lstp_A -invariant.

Definition 5.23. Let $I = (a_i : i < \omega)$ be a sequence of tuples. We say that it is a *Morley sequence* over A if there is some global type p , non-forking over A such that $I \models p^{(\omega)}|_A$.

Lemma 5.24. Let $\bar{b} = (b_i : i < \omega)$ be an A -indiscernible sequence such that $p = \text{tp}(\bar{b}/A)$ does not fork over A . Then there is some global non-forking extension $q(x_0, x_1, \dots)$ of p which is the type of some \mathcal{U} -indiscernible sequence.

Proof. Let p' be any non-forking extension of $\text{tp}(\bar{b}/A)$. Let $(d_i : i < \omega) \models p'$. We can find some \mathcal{U} -indiscernible sequence $(e_i : i < \omega)$ which is based on the sequence $(d_i : i < \omega)$ over \mathcal{U} . Let $q = \text{tp}(e_i : i < \omega)$. Then q does not fork over A and its restriction to A is p . \square

Lemma 5.25. Let A be an extension base, and $\phi(x; b) \in L(\mathcal{U})$ divide over A . Then there is a Morley sequence $(b_i : i < \omega)$ over A witnessing dividing (i.e., $b_0 = b$ and the conjunction $\bigwedge_{i < \omega} \phi(x; b_i)$ is inconsistent).

Proof. Let $(b'_i : i < \omega)$ be some A -indiscernible sequence with $b'_0 = b$ and such that $\{\phi(x; b'_i)\}$ is k -inconsistent. Let $p = \text{tp}((b'_i)_{i < \omega}/A)$. By assumption on A , p does not fork over A , so by the previous lemma, there is some $q(x_0, \dots) \in S(\mathcal{U})$ a non-forking extension of p and some $r(x) \in S(\mathcal{U})$ such that the type q restricted to any one of its variable is equal to r . Let $(\bar{b}^t : t < \omega)$ be a Morley sequence of q over A , where $\bar{b}^t = (b_i^t : i < \omega)$.

For each $t < \omega$, the type $\{\phi(x; b_i^t) : i < \omega\}$ is k -inconsistent. By NTP_2 , there is some $\eta : \omega \rightarrow \omega$ such that $\bigwedge_{t < \omega} b_{\eta(t)}^t$ is consistent. As the sequence $(b_{\eta(t)}^t : t < \omega)$ is a Morley sequence of r over A , we have what we want. \square

Lemma 5.26 (Broom lemma). Let $A \subset \mathcal{U}$ be an extension base and $\pi(x)$ be a partial type over \mathcal{U} which is Lstp_A -invariant. Assume that for some formulas $\psi(x; b)$ and $\phi_i(x; c)$, $i < n$ in $L(\mathcal{U})$ we have

$$\pi(x) \vdash \psi(x; b) \vee \bigvee_{i < n} \phi_i(x; c),$$

where $b \perp_A c$ and for each $i < n$, $\phi_i(x; c)$ divides over A . Then

$$\pi(x) \vdash \psi(x; b).$$

Proof. We prove the result by induction on n . The case $n = 0$ is trivial.

Assume we know it for n and let $\pi(x) \vdash \psi(x; b) \vee \bigvee_{i < n+1} \phi_i(x; c)$. By the previous lemma, there is a Morley sequence $(c_i : i < \omega)$, $c_0 = c$ such that $\{\phi(x; c_i) : i < \omega\}$ is k -inconsistent. Conjugating by an automorphism, we may assume that $b \perp_A (c_i)_{i < \omega}$, thus the sequence $(c_i)_{i < \omega}$ is indiscernible over Ab (since $\text{tp}(b/A(c_i))$ is Lstp_A -invariant).

By Lstp_A -invariance of $\pi(x)$, we have $\pi(x) \vdash \psi(x; b) \vee \bigvee_{i < n+1} \phi_i(x; c_j)$ for all $j < \omega$. In particular, we have

$$\pi(x) \vdash \psi(x; b) \vee \bigwedge_{j < k} \bigvee_{i < n+1} \phi_i(x; c_j).$$

By assumption on k , this implies

$$\pi(x) \vdash \psi(x; b) \vee \bigvee_{\substack{j < k \\ i < n}} \phi_i(x; c_j).$$

For any $j < k$, we have $b \perp_A c_{\geq j}$, thus $b \perp_{Ac_j} c_{> j}$. Also we have $c_{> j} \perp_A c_j$. Therefore by transitivity (Lemma 5.10) $bc_{> j} \perp_A c_j$.

By the induction hypothesis, and as

$$\pi(x) \vdash \left[\psi(x; b) \vee \bigvee_{\substack{0 \leq j < k \\ i < n}} \phi_i(x; c_j) \right] \vee \bigvee_{i < n} \phi_i(x; c_j)$$

we have

$$\pi(x) \vdash \psi(x; b) \vee \bigvee_{\substack{0 \leq j < k \\ i < n}} \phi_i(x; c_j).$$

Iterating, we obtain $\pi(x) \vdash \psi(x; b)$. □

Remark 5.27. If in the hypothesis of the lemma, we remove the assumption that $b \perp_A c$, then the conclusion becomes

$$\pi(x) \vdash \bigvee_{i < m} \psi(x; b_i),$$

for some b_i 's satisfying $\text{Lstp}(b_i/A) = \text{Lstp}(b/A)$.

Proof. Let $\pi(x)$, $\psi(x; b)$ and $\phi^i(x; c)$ be as in the hypothesis of the lemma, except that we do not assume $b \perp_A c$. Consider the partial type $\pi'(x) = \pi(x) \cup \{\neg\psi(x; b') : \text{Lstp}(b'/A) = \text{Lstp}(b/A)\}$. Then $\pi'(x)$ is Lstp_A -invariant and by hypothesis on $\pi(x)$, we have $\pi'(x) \vdash x \neq x \vee \bigvee \phi^i(x; c)$. The previous lemma applies to give $\pi'(x) \vdash x \neq x$. Therefore $\pi'(x)$ is inconsistent and the conclusion follows by compactness. \square

Corollary 5.28. *Let $M \models T$ and $p \in S(M)$. Then p has a non-forking heir over M .*

Proof. Assume not. Then the partial type $\pi(x) = p(x) \cup \{\psi(x; b) : b \in \mathcal{U}, p \vdash \psi(x; b'), \forall b' \in M\} \cup \{\phi(x; c) : c \in \mathcal{U}, \neg\phi(x; c) \text{ divides over } M\}$ is inconsistent. By compactness, we find a formula $\psi(x; b)$ from the first set and finitely many formulas $\phi^i(x; c)$, $i < n$ from the second set such that $p(x) \vdash \neg\psi(x; b) \vee \neg \bigvee_{i < n} \phi^i(x; c)$. By Remark 5.27, we have $p(x) \vdash \bigvee_{i < m} \neg\psi(x; b_i)$, where $b_i \equiv_M b$. As p is a type over M , by compactness, there is some formula $\theta(x_0, \dots, x_{m-1})$ such that

$$p(x) \vdash \forall y_0, \dots, y_{m-1} \theta(\bar{y}) \rightarrow \bigvee_{i < m} \neg\psi(x; y_i).$$

Taking such y_0, \dots, y_{m-1} in M contradicts the choice of formulas $\psi(x; y)$. \square

5.5 Strict non-forking

Definition 5.29. Let $A \subseteq B$ and $a \in \mathcal{U}$. We say that $p = \text{tp}(a/B)$ is *strictly non-forking* over A if there is a global extension p' of p such that p' is non-forking over A and such that for all $C \subseteq \mathcal{U}$, if $a_0 \models p'|_C$, then $C \perp_A a_0$.

We will write $a \perp_A^{st} b$ to mean that $\text{tp}(a/Ab)$ is strictly non-forking over A .

Remark 5.30. In particular, if $\text{tp}(a/B)$ is strictly non-forking over A , then we have $a \perp_A B$ and $B \perp_A a$. It is an open question whether the converse holds (assuming that A is an extension base).

EXAMPLE 5.31. Let $p \in S(M)$, then p is strictly non-forking over M . This is witnessed by any global non-forking heir p' of p (using Corollary 5.28).

We now generalize this to types over arbitrary extension bases.

Proposition 5.32. *Let A be an extension base and $p \in S(A)$. Then p is strictly non-forking over A .*

Proof. Let $a \models p$. The proof is essentially the same as that of Corollary 5.28. Namely, we consider the partial type $p(x) \cup \{\psi(x; b) : b \in \mathcal{U}, \neg\psi(a; y) \text{ divides over } A\} \cup \{\phi(x; c) : c \in \mathcal{U}, \neg\phi(x; c) \text{ divides over } A\}$. If this type is inconsistent, then we can find some $\psi(x; b)$ from the first set and finitely many $\phi^i(x; c)$, $i < n$ from the second set such that

$$p(x) \vdash \neg\psi(x; b) \vee \bigvee_{i < n} \neg\phi^i(x; c).$$

By Remark 5.27, we have $p(x) \vdash \bigvee_{i < m} \neg\psi(x; b_i)$ where $\text{Lstp}(b_i/A) = \text{Lstp}(b/A)$. Let $\bar{b} = b_0 \wedge \dots \wedge b_{m-1}$ and take $q \in S(\mathcal{U})$ a non-forking extension of $\text{tp}(\bar{b}/A)$ (as A is an extension base). If $\bar{b}' \models q|_{Aa}$, then by hypothesis on $\psi(x; y)$, $\psi(a; b_i)$ holds for all i . A contradiction. \square

Definition 5.33. The sequence $(a_i : i < \alpha)$ is a *strict non-forking sequence* over A , if for any $i < \alpha$, we have $a_i \downarrow_A^{st} a_{<i}$.

Lemma 5.34. *Let A be an extension base. Let $(a_i : i < \alpha)$ be a strict non-forking sequence over A and $(I_i : i < \alpha)$ a sequence of A -indiscernible sequences such that I_i begins with a_i . Then we can find sequences $(J_i : i < \alpha)$ mutually indiscernible over A such that $J_i \equiv_{Aa_i} I_i$ for each $i < \alpha$.*

Proof. Without loss, the sequences I_i 's are indexed by ω . We may assume that $\alpha = n < \omega$ and we show the result by induction on n . Assume it is true for n and let $(a_i : i < n+1)$ and $(I_i : i < n+1)$ be given. By the induction hypothesis, we may assume that the sequences $\{I_i : i < n\}$ are mutually indiscernible over A . We have $a_n \downarrow_A^{st} a_{<n}$. Conjugating by an automorphism fixing A and the sequence $(a_i : i < n+1)$, we may assume that $a_n \downarrow_A^{st} a_{<n} I_{<n}$. This implies:

- (1) the sequences $\{I_i : i < n\}$ are mutually indiscernible over Aa_n ;
- (2) $a_{<n} I_{<n} \downarrow_A^{st} a_n$.

By Lemma 5.9, $\text{tp}(a_{<n} I_{<n}/Aa_n)$ has an extension q over AI_n such that I_n is indiscernible over a realization of q . Therefore moving I_n by an automorphism over Aa_n , we may assume that I_n is indiscernible over $Aa_{<n} I_{<n}$.

For $k < n+1$, write $I_k = (a_i^k : i < \omega)$. Let $\eta : n+1 \rightarrow \omega$ be any function. As I_n is indiscernible over $AI_{<n}$, we have

$$a_{\eta(0)}^0 \cdots a_{\eta(n)}^n \equiv_A a_{\eta(0)}^0 \cdots a_{\eta(n-1)}^{n-1} a_n.$$

Then by (1) above,

$$a_{\eta(0)}^0 \cdots a_{\eta(n)}^n \equiv_A a_0 \cdots a_n.$$

By Ramsey's theorem, we can find sequences $(I'_k : k < n)$ mutually indiscernible over AI_n and based on $(I_k : k < n)$. Write $I'_k = (a_i^k : i < \omega)$. Then by the previous observation we have $a_0^0 \cdots a_0^{n-1} \equiv_{Aa_n} a_0 \cdots a_{n-1}$. Therefore we may assume that $a_0^k = a_k$ for all $k < n$.

Then the sequences $I'_0, \dots, I'_{n-1}, I_n$ have the required properties. \square

Proposition 5.35. *Let A be an extension base, and let $(a_i : i < \kappa)$ be a strict non-forking sequence over A . Let b such that $\text{dp-rk}(b/A) < \kappa$. Then there is $i < \kappa$ such that $\text{tp}(b/Aa_i)$ does not divide over A .*

Proof. Assume that $\text{tp}(b/Aa_i)$ divides over A for all $i < \kappa$. Then, for each $i < \kappa$ we can find some $\phi_i(x; a_i) \in \text{tp}(b/Aa_i)$ and an A -indiscernible sequence $I_i = (a_j^i : j < \omega)$ such that $a_0^i = a_i$ and $\bigwedge_{j < \omega} \phi_i(x; a_j^i)$ is inconsistent. By the previous proposition, we may assume that the sequences I_i are mutually indiscernible over A . Then for $i < \kappa$, I_i is not indiscernible over Ab , which contradicts the definition of dp-rank. \square

We will show in Theorem 5.37 that forking equals dividing over extension bases, therefore one can replace “does not divide” by “does not fork” in the previous proposition.

Proposition 5.36. *Let A be an extension base and $(a_i : i < \omega)$ indiscernible and strict non-forking over A . Assume that the formula $\phi(x; a_0)$ divides over A , then the partial type $\{\phi(x; a_i) : i < \omega\}$ is inconsistent.*

We say the strict non-forking sequences *witness dividing*.

Proof. Assume that $\{\phi(x; a_i) : i < \omega\}$ is consistent. Then we may increase the sequence to some A -indiscernible $(a_i : i < |T|^+)$, strict non-forking over A such that there is $b \models \bigwedge_{i < |T|^+} \phi(x; a_i)$. Then $\text{tp}(b/Aa_i)$ divides over a_i for each i , and this contradicts Proposition 5.35. \square

Theorem 5.37. *Let A be an extension base and $\phi(x; a)$ a formula, then $\phi(x; a)$ forks over A if and only if it divides over A .*

Proof. If $\phi(x; a)$ divides over A , then it forks over A . Conversely, assume that $\phi(x; a)$ forks over A . Then there are formulas $\psi_i(x; c) \in L(\mathcal{U})$, $i < n$, such that $\phi(x; a) \vdash \bigvee_{i < n} \psi_i(x; c)$ and each $\psi_i(x; c)$ divides over A . Let $q \in S(\mathcal{U})$ be a strict non-forking extension of $\text{tp}(ac/A)$ (using Proposition 5.32) and let $(a_i c_i : i < \omega)$ be a Morley sequence of q over A . If $\phi(x; a)$ does not divide over A , then there is $b \models \bigwedge_{i < \omega} \phi(x; a_i)$. Then trimming the sequence if necessary, we may assume without loss that $b \models \bigwedge_{i < \omega} \psi_0(x; c_i)$. This contradicts Proposition 5.36. \square

Lowness

Let A be any set of parameters and let $\phi(x; y)$ be a formula over A . Let k be the alternation number of $\phi(x; y)$, that is the maximal k for which we can find an A -indiscernible sequence $(b_i : i < \omega)$ such that $\bigwedge_{i < k} \neg(\phi(x; b_i) \leftrightarrow \phi(x; b_{i+1}))$ is consistent. Let now $(b_i : i < \omega)$ be any A -indiscernible sequence. Assume that $\bigwedge_{i < k+1} \phi(x; b_i)$ is consistent and we show that $\bigwedge_{i < \omega} \phi(x; b_i)$ is consistent.

To see this, increase the sequence to an A -indiscernible $(b_i : i < \omega \cdot (k+1))$. By indiscernability, $\bigwedge_{i < k+1} \phi(x; b_{\omega \cdot i})$ is consistent and let a realize it. As k is the alternation number, there is some $i < k+1$ such that a realizes $\bigwedge_{j < \omega} \phi(x; b_{\omega \cdot i + j})$, otherwise the truth value of $\phi(a; b_i)$ would alternate too much on the sequence. Thus again by indiscernability, $\bigwedge_{i < \omega} \phi(x; b_i)$ is consistent.

Proposition 5.38. *Let A be an extension base and $\phi(x; y)$ a formula over A . Then the set of b 's such that $\phi(x; b)$ forks over A is type-definable over A .*

Proof. By the paragraph before the proof (and Theorem 5.37), we know that $\phi(x; b)$ forks over A , if and only if it divides over A , if and only if there is some A -indiscernible sequence $(b_i : i < \omega)$ such that:

- $b_0 = b$;
- $\bigwedge_{i < \text{alt}(\phi)+1} \phi(x; b_i)$ is inconsistent.

This is easily seen to be a type-definable condition on b . \square

Theories satisfying the conclusion of the previous proposition are called *low* theories. Thus NIP theories are low.

Historical Remarks

The definition and general properties of forking and dividing are due to Shelah. See for example his book [81]. Properties of forking specific to NIP theories

presented in Section 5.2 are proved in Shelah's [83] and Hrushovski and Pillay's [47]. The proof that forking equals dividing is from Chernikov and Kaplan's [21]. We follow closely the simplified account given by Adler in [1]. The definition and properties of strict non-forking appeared in [83] with slightly erroneous statements and proofs. It is investigated in [54].

Chapter 6

Finite combinatorics

In this chapter, we introduce tools from finite combinatorics and probability theory.

It was first observed by Laskowski in [57] that the notion of NIP formula had an analogue in combinatorics under the name *finite VC-dimension*. More precisely, a formula $\phi(x; y)$ is NIP if and only if the class $\{\phi(M; b) : b \in M\}$ of subsets of $M^{|x|}$ has finite VC-dimension. This is a central notion in machine learning theory, as classes of finite VC-dimension coincide with *learnable classes*. See [30]. This theory builds on the theorem of Vapnik and Chervonenkis (from which the name *VC* originates) which states a uniform law of large numbers for such classes.

We give a proof of that theorem and of the so-called (p, q) -theorem of Alon-Kleitman and Matousek. The later could be restated in model-theoretic terms, but no proof is known using the classical tools of first order logic. We will therefore change the framework in this chapter and consider a purely combinatorial setting. We will come back to first order structures at the end to prove uniformity of honest definitions.

6.1 VC-dimension

We consider the following situation: X is a set (finite or infinite), and \mathcal{S} is a family of subsets of X . Such a pair (X, \mathcal{S}) is called a *set system*. For most purposes, we can forget about the base set X , or in other words, take X to be $\bigcup \mathcal{S}$.

Let $A_0 \subset X$. We say that the family \mathcal{S} *shatters* A_0 , if for every $A \subseteq A_0$,

there is a set S in \mathcal{S} such that $S \cap A_0 = A$.

The family \mathcal{S} has VC-dimension at most n (written $\text{VC}(\mathcal{S}) \leq n$), if there is no $A_0 \subseteq X$ of cardinality $n+1$ such that \mathcal{S} shatters A_0 . We say that \mathcal{S} is of VC-dimension n if it is of VC-dimension at most n and shatters some subset of size n .

If for each n we can find a subset of X of cardinality n shattered by \mathcal{S} , then we say that \mathcal{S} has infinite VC-dimension (and write $\text{VC}(\mathcal{S}) = \infty$).

For a more careful analysis, we define the *shatter function* $\pi_{\mathcal{S}}$ from \mathbb{N} to \mathbb{N} as follows: $\pi_{\mathcal{S}}(n)$ is the maximum over all $A_0 \subseteq X$ of cardinality n of $|\{C \cap A_0 : C \in \mathcal{S}\}|$. Note that $\pi_{\mathcal{S}}$ is bounded by 2^n , and we have:

$$\text{VC}(\mathcal{S}) < n \iff \pi_{\mathcal{S}}(n) < 2^n.$$

Given a set system (X, \mathcal{S}) , we define the *dual* set system as the set system (X^*, \mathcal{S}^*) , where $X^* = \mathcal{S}$ and $\mathcal{S}^* = \{\mathcal{S}_a : a \in X\}$, with $\mathcal{S}_a = \{S \in \mathcal{S} : a \in S\}$. We then define the *VC codimension* of \mathcal{S} (written $\text{VC}^*(\mathcal{S})$) as the VC-dimension of \mathcal{S}^* . We also define the dual shatter function $\pi_{\mathcal{S}}^*$ as the shatter function of \mathcal{S}^* .

EXAMPLE 6.1. Let $\phi(x; y)$ be a formula and fix a model M . Then $\phi(x; y)$ is NIP if and only if the class $\mathcal{S}_{\phi}(M) = \{\phi(M; b) : b \in M\}$ is of finite VC-dimension. The VC-dimension of $\phi(x; y)$ as defined in Section 2.1 coincides with VC-dimension of that class. In particular, it is independent of the choice of M . We similarly define π_{ϕ} and the dual objects. Note that the VC codimension of $\phi(x; y)$ is the VC-dimension of $\phi^*(y; x)$ where $\phi^*(y; x) = \phi(x; y)$.

Lemma 6.2. *We have $\text{VC}^*(\mathcal{S}) < 2^{\text{VC}(\mathcal{S})+1}$ and $\text{VC}(\mathcal{S}) < 2^{\text{VC}^*(\mathcal{S})+1}$.*

In particular \mathcal{S} has finite VC-dimension, if and only if it has finite VC-codimension.

Proof. Assume $\text{VC}(\mathcal{S}) \geq 2^n$. Then there is some subset $A \subseteq X$ of size 2^n shattered by \mathcal{S} . Write $A = \{a_C : C \subseteq [n]\}$. For each $k < n$, let $S_k \in \mathcal{S}$ be such that $S_k \cap A = \{a_C : \{k\} \subseteq C \subseteq [n]\}$. Then easily, the family $\{S_k : k \leq n\}$ is shattered by \mathcal{S}^* . It follows that $\text{VC}^*(\mathcal{S}) \geq n$. This proves the second inequality. The first one is proved similarly. \square

The following lemma states that the shatter function $\pi_{\mathcal{S}}(n)$ is either always equal to 2^n , or has polynomial growth.

Lemma 6.3 (Sauer's lemma). *Let \mathcal{S} be a class of VC-dimension at most k . Then, for $n \geq k$, we have $\pi_{\mathcal{S}}(n) \leq \sum_{k=0}^n \binom{n}{k}$. In particular $\pi_{\mathcal{S}}(n) = O(n^k)$.*

Proof. First notice that the bound is tight: take \mathcal{S} to be the family of subsets of X of cardinality k . Then \mathcal{S} has VC-dimension exactly k and we see that its shatter function is equal to the bound in the statement of the lemma. The idea of the proof is to reduce the situation to this case by modifying the elements of \mathcal{S} , making them as small as possible without changing the cardinality of the family nor its VC-dimension.

Fix an integer $n \geq k$. If \mathcal{S} contradicts the bound, then this is also true of some finite subfamily of \mathcal{S} , so without loss we may assume that \mathcal{S} is finite. Similarly, we may assume that X is finite of cardinality n . Say $X = \{x_1, \dots, x_n\}$.

We define iteratively families $\mathcal{S}_0, \dots, \mathcal{S}_n$. Set $\mathcal{S}_0 = \mathcal{S}$.

Let $k < n$, and assume that \mathcal{S}_k has been defined. Go through the sets in \mathcal{S}_k one by one. For each $S \in \mathcal{S}_k$, if $x_k \in S$, and if $S \setminus \{x_k\}$ is not a set in \mathcal{S}_k , replace S by $S \setminus \{x_k\}$. Let \mathcal{S}_{k+1} be the resulting family.

The following three facts can be easily checked by induction on k :

- (i) for each k , the cardinality of \mathcal{S}_k is the same as that of \mathcal{S} ;
- (ii) let $S \in \mathcal{S}_k$ and $A = S \cap \{x_1, \dots, x_k\}$; then for every $A_0 \subseteq A$, the set $A_0 \cup (S \setminus A)$ is in \mathcal{S}_k ;
- (iii) any $A \subseteq X$ shattered by \mathcal{S}_k is also shattered by \mathcal{S}_{k-1} .

Fact (ii) implies in particular that if $S \in \mathcal{S}_n$, then S is shattered by \mathcal{S}_n . From (iii), it follows that the VC-dimension of \mathcal{S}_n is not greater than that of \mathcal{S} . Therefore no set in \mathcal{S}_n can have cardinality greater than k . As $|\mathcal{S}_n| = |\mathcal{S}|$, we obtain the required bound. \square

We now define the *VC density* of \mathcal{S} to be $\text{vc}(\mathcal{S}) = \limsup_{n \rightarrow \infty} \frac{\log(\pi_{\mathcal{S}}(n))}{\log n}$. In other words, $\text{vc}(\mathcal{S})$ is the smallest $r \geq 0$ for which we have $\pi_{\mathcal{S}}(n) = O(n^r)$. Similarly, we define the VC-codensity, as $\text{vc}^*(\mathcal{S}) = \text{vc}(\mathcal{S}^*)$.

By the previous result, we always have $\text{vc}(\mathcal{S}) \leq \text{VC}(\mathcal{S})$. For many purposes, the VC density is a more appropriate notion than the VC-dimension.

We now state and prove the fundamental theorem of Vapnik and Chervonenkis. It is a uniform version of the law of large numbers for set systems of finite VC-dimension.

First, we fix a notation. For $S \in \mathcal{S}$ and $(x_1, \dots, x_n) \in X^n$, we define $\text{Av}(x_1, \dots, x_n; S)$ as being equal to $\frac{1}{n}|S \cap \{x_1, \dots, x_n\}|$. It is the measure

of S estimated on the finite set $\{x_1, \dots, x_n\}$. The weak law of large numbers (Proposition C.3) states that for fixed $S \in \mathcal{S}$, and $\epsilon > 0$, we have $\mu^n(|\text{Av}(x_1, \dots, x_n; S) - \mu(S)| \geq \epsilon) \rightarrow 0$. It follows from its proof, that the speed of convergence is bounded by a function depending only on ϵ . In other words, with high probability, sampling on a tuple (x_1, \dots, x_n) selected at random gives a good estimate of the measure of S . The VC-theorem states that if \mathcal{S} is of finite VC-dimension, then sampling on a random tuple (x_1, \dots, x_n) gives a good estimate of the measures of all the sets in \mathcal{S} .

Theorem 6.4 (VC-theorem). *Let (X, μ) be a finite probability space, and $\mathcal{S} \subseteq \mathfrak{P}(X)$ a family of subsets, then for $\epsilon > 0$ we have:*

$$\mu^n \left(\sup_{S \in \mathcal{S}} |\text{Av}(x_1, \dots, x_n; S) - \mu(S)| > \epsilon \right) \leq 4\pi_{\mathcal{S}}(2n) \exp \left(-\frac{n\epsilon^2}{8} \right).$$

Proof. For $\bar{x} = (x_1, \dots, x_n)$, $\bar{x}' = (x'_1, \dots, x'_n)$ and $S \in \mathcal{S}$, we let $\delta(\bar{x}, \bar{x}'; S)$ be equal to $|\text{Av}(x_1, \dots, x_n; S) - \text{Av}(x'_1, \dots, x'_n; S)|$. Let also $\delta_i(\bar{x}, \bar{x}'; S) = \frac{1}{n}(\mathbf{1}_S(x_i) - \mathbf{1}_S(x'_i))$. Thus $\delta(\bar{x}, \bar{x}'; S) = |\sum_{i=1..n} \delta_i(\bar{x}, \bar{x}'; S)|$.

Claim: We have:

$$\mu^{2n} \left(\sup_{S \in \mathcal{S}} \delta(\bar{x}, \bar{x}'; S) > \epsilon/2 \right) \leq 2\pi_{\mathcal{S}}(2n) \exp \left(-\frac{n\epsilon^2}{8} \right).$$

Proof: Fix some integer n and some tuples $\bar{x}, \bar{x}' \in X^n$. Fix also some $S \in \mathcal{S}$. For $i \leq n$, we let \mathbb{Z}_2 act on (x_i, x'_i) by flipping the two variables, and we let $(\mathbb{Z}_2)^n$ act on $\bar{x} \bar{x}'$ by the product action. Let $\Omega_S = \{i \in \{1, \dots, n\} : \delta_i(\bar{x}, \bar{x}'; S) \neq 0\}$.

Let Y be the probability space $(\mathbb{Z}_2)^n$ equipped with the normalized counting measure. We consider the random variable $f : Y \rightarrow \mathbb{R}$ defined by $f(\sigma) = \delta(\sigma(\bar{x}, \bar{x}'); S)$ and for each i , we let f_i be the random variable defined by $f_i(\sigma) = \delta_i(\sigma(\bar{x}, \bar{x}'); S)$. Then $f = |\sum f_i|$. In fact, for $i \notin \Omega_S$, f_i is uniformly equal to zero. Thus we have $f = |\sum_{i \in \Omega_S} f_i|$.

For $i \in \Omega_S$, the random variable f_i satisfies $\text{Prob}(f_i = -1/n) = \text{Prob}(f_i = 1/n) = 1/2$. Also, the variables f_i are independent. Let $A_S \subset Y$ be the set of $\sigma \in Y$ for which $f(\sigma) \geq \epsilon/2$. By Chernoff's bound (Proposition C.4), we have

$$|A_S| \leq 2^{|\Omega_S|+1} \exp \left(-\frac{|\Omega_S|\epsilon^2}{8} \right) \leq 2^{n+1} \exp \left(-\frac{n\epsilon^2}{8} \right).$$

The set A_S depends only on $S \cap \{x_1, \dots, x_n, x'_1, \dots, x'_n\}$. As S varies in \mathcal{S} , there are at most $\pi_{\mathcal{S}}(2n)$ values for the sets A_S . Thus $A_* = \bigcup_{S \in \mathcal{S}} A_S$ has measure at

most $\alpha := 2\pi_{\mathcal{S}}(2n) \exp(-n\epsilon^2/8)$. For $\sigma \notin A_*$, we have $\sup_{S \in \mathcal{S}} \delta(\sigma(\bar{x}, \bar{x}'); S) \leq \epsilon/2$.

Consider the space $X^{2n} \times Y$ equipped with the product measure. Define $D \subset X^{2n} \times Y$ by $(\bar{x}, \bar{x}', \sigma) \in D$ if $\sup_{S \in \mathcal{S}} \delta(\sigma(\bar{x}, \bar{x}'); S) > \epsilon/2$. Then for fixed $\bar{x}, \bar{x}' \in X^{2n}$, the fiber $D_{(\bar{x}, \bar{x}')} \subseteq Y$ has measure at most α . Therefore the measure of D is at most α . On the other hand, given $\sigma, \sigma' \in Y$, the fibers $D^{(\sigma)}, D^{(\sigma')}$ have the same measure because the measure on X^{2n} is symmetric. Therefore that measure is at most α , and the claim follows.

Now, we may assume that $n > 2/\epsilon^2$, since otherwise the inequality is trivial. Let $X_0 \subseteq X^n$ be the set of $\bar{a} \in X^n$ such that $\mu^n(\sup_{S \in \mathcal{S}} \delta(\bar{x}, \bar{a}; S) > \epsilon/2) \geq 1/2$. We have $\mu^n(X_0) \leq 4\pi_{\mathcal{S}}(2n) \exp\left(-\frac{n\epsilon^2}{8}\right)$. Pick $\bar{a} \in X^n \setminus X_0$ and some $S \in \mathcal{S}$. By Chebyshev's inequality, we have

$$\text{Prob}(|\text{Av}(x_1, \dots, x_n; S) - \mu(S)| > \epsilon/2) \leq 1/(n\epsilon^2) < 1/2,$$

(see the proof of Proposition C.3). It follows that there is $\bar{x} \in X^n$ satisfying both:

- $\delta(\bar{a}, \bar{x}; S) \leq \epsilon/2$;
- $|\text{Av}(\bar{x}; S) - \mu(S)| \leq \epsilon/2$.

Remembering the definition of δ , this implies that $|\text{Av}(\bar{a}; S) - \mu(S)| \leq \epsilon$. This finishes the proof. \square

Remark 6.5. In fact the theorem is still true if X is infinite, but we have to add the following measurability assumptions:

- each set $S \in \mathcal{S}$ is measurable;
- for each n , the function $X^n \rightarrow \mathbb{R}$, $(x_1, \dots, x_n) \mapsto \sup_{S \in \mathcal{S}} |\text{Av}(x_1, \dots, x_n) - \mu(S)|$ is measurable;
- for each n , the function $X^{2n} \rightarrow \mathbb{R}$, $(x_1, \dots, x_n, x'_1, \dots, x'_n) \mapsto \sup_{S \in \mathcal{S}} \delta(\bar{x}, \bar{x}'; S)$ is measurable.

The first condition implies the other two when the family \mathcal{S} is countable, and of course they always hold when X is finite. The proof then goes through.

To see how the second and third hypothesis might fail, consider the case of $X = \omega_1$. Let \mathcal{B} be the σ -algebra generated by the intervals. Let μ be defined on \mathcal{B} by $\mu(A) = 1$ if A contains an end segment of X and $\mu(A) = 0$ otherwise. Note that this defines a σ -additive measure on (X, \mathcal{B}) . Let \mathcal{S} to be the family of intervals of X . It has VC-dimension 2. We leave it to the reader to check

that the VC-theorem does not hold for \mathcal{S} . (In view of Corollary 6.6 it is enough to check that there are no ϵ -nets, for $\epsilon < 1$).

If (X, \mathcal{S}) is a set system, an ϵ -net for \mathcal{S} is a finite subset $X_0 = \{x_1, \dots, x_n\}$ of X such that for all $S \in \mathcal{S}$, we have

$$\left| \frac{1}{n} |S \cap \{x_1, \dots, x_n\}| - \mu(S) \right| \leq \epsilon.$$

Corollary 6.6. *Let $k > 0$ and $\epsilon > 0$, then there is N such that any set system \mathcal{S} on a finite probability space (X, μ) with $VC(\mathcal{S}) \leq k$ admits an ϵ -net of size at most N .*

Proof. By Sauer's lemma, we know that there is some polynomial P_k depending only on k such that $\pi_{\mathcal{S}}(n) \leq P_k(n)$. Take N such that $4P_k(2N) \exp\left(-\frac{N\epsilon^2}{8}\right) < 1$ and apply the VC-theorem 6.4. \square

6.2 The (p, q) -theorem

Let $p \geq q$ be two integers. A set system (X, \mathcal{S}) has the (p, q) -property if for every $X_1 \subseteq X$ of cardinality $\geq p$, there is $S \in \mathcal{S}$ such that $|X_1 \cap S| \geq q$.

The following theorem will be an important ingredient in the proof of uniformity for honest definitions.

Theorem 6.7. *Let $p \geq q$ be two integers. Then there is an integer N such that the following holds:*

For any set system (X, \mathcal{S}) , assume:

- $X = \cup \mathcal{S}$ is finite;*
- $VC(\mathcal{S}) \leq q$;*
- (X, \mathcal{S}) has the (p, q) -property.*

Then we can find $S_1, \dots, S_N \in \mathcal{S}$ such that $\bigcup_{i=1..N} S_i = X$.

Actually, we will only need a special case of this theorem with $p = q$ big enough with respect to $VC(\mathcal{S})$. We state this as a separate corollary. We will only prove this corollary (without using the theorem) and refer the reader to the references for the proof of the full theorem.

Corollary 6.8. *Let $k \in \mathbb{N}$, then there are two integers q and N such that for every finite X and $\mathcal{S} \subseteq \mathfrak{P}(X)$ a family of VC-dimension at most k , if for every $X_0 \subset X$ of size $\leq q$, we can find $S \in \mathcal{S}$ containing X_0 , then there are $S_1, \dots, S_N \in \mathcal{S}$ whose union is the whole of X .*

Proof. (Of the corollary)

Let $\epsilon = 1/3$.

By Corollary 6.6, there is some q such that for every set system \mathcal{S} of VC-dimension $\leq k$ on a finite set X and any probability measure μ on X , there are $x_1, \dots, x_q \in X$ such that for any $S \in \mathcal{S}$,

$$\left| \mu(S) - \frac{|\{i : x_i \in S\}|}{n} \right| \leq \epsilon.$$

Let (X, \mathcal{S}) be a set system of VC-dimension $\leq k$ with X finite and having the (q, q) -property. Then it follows that for any probability measure μ on X , we can find some $S \in \mathcal{S}$ with $\mu(S) \geq 1 - \epsilon$.

We now need the following result, known as Farkas' lemma. We refer the reader to any introductory text on convex analysis or linear programming for a proof. See for example [76, page 90].

Fact 6.9 (Farkas' lemma). *Let A be a matrix in $\mathcal{M}_{m,n}(\mathbb{R})$ and $b \in \mathbb{R}^n$, then the following are equivalent:*

- (i) $\exists x \in \mathbb{R}_+^n$ such that $Ax \leq b$;
- (ii) for all $y \in \mathcal{M}_{1,m}(\mathbb{R}_+)$, $yA \geq 0$ implies $yb \geq 0$.

Write $X = \{x_1, \dots, x_{m-1}\}$ and $\mathcal{S} = \{S_1, \dots, S_n\}$. We define a matrix $A \in \mathcal{M}_{n,m}$ by setting $A_{m,j} = 1$ for all $j \leq n$, and for $i \leq m-1$ and $j \leq n$, we set $A_{i,j} = -1$ if $x_i \in S_j$ and 0 otherwise. We also define $b \in \mathbb{R}^m$ by $b_i = -(1-\epsilon)$ for $i \leq m-1$ and $b_m = 1$.

We check that condition (ii) of Farkas' lemma is satisfied. So let $y = (a_1, \dots, a_{m-1}, a_m) \in \mathcal{M}_{1,m}(\mathbb{R}_+)$ be such that $yA \geq 0$. This means that for any $j \leq n$, we have $\sum_{i=1}^m a_i A_{i,j} \geq 0$. Hence, for any $j \leq n$, $\sum_{\{i: x_i \in S_j\}} a_i \leq a_m$.

Let $a_* = \sum_{i \leq m-1} a_i$. Let $\tilde{\mu}$ be the measure on X defined by giving weight a_i/a_* to the point x_i , for all $i \leq m-1$. Then the previous inequality implies that no set S_j has measure greater than a_m/a_* . On the other hand, by what we have shown above, there is some $S \in \mathcal{S}$ such that $\tilde{\mu}(S) \geq (1-\epsilon)$. It follows that $a_m/a_* \geq 1 - \epsilon$. And this gives exactly $yb \geq 0$.

By Farkas' lemma, we conclude that (i) above holds, namely that there is some $d = (d_1, \dots, d_n)^T$, with each $d_i \geq 0$ such that $Ad \leq b$. Decoding, this

gives that for $i \leq m-1$, $\sum_{j: x_i \in S_j} d_j \geq 1 - \epsilon$ and $\sum_{j=1}^m d_j \leq 1$. Increasing d_0 if necessary, we may assume that actually $\sum_{j=1}^m d_j = 1$.

We now consider the dual set system: to $x \in X$, we associate $E_x = \{S \in \mathcal{S} : x \in S\}$. Let $\mathcal{E} = \{E_x : x \in X\}$. Then $\text{VC}(\mathcal{E}) \leq 2^{k+1}$. Equip the set \mathcal{S} by the measure μ^* defined by the weights d_i . By Corollary 6.6 again, we can find N depending only on k and ϵ and $S'_1, \dots, S'_l \in \mathcal{S}$ such that for all $x \in X$,

$$\left| \mu^*(E_x) - \frac{|\{l : x \in S'_l\}|}{N} \right| \leq \epsilon.$$

For any $x \in X$, we have $\mu^*(E_x) = \sum_{\{j: x \in S_j\}} d_j \geq 1 - \epsilon > \epsilon$, thus $|\{l : x \in S'_l\}| > 0$. It follows that $\bigcup_{l=1}^N S'_l = X$. \square

6.3 Uniformity of honest definitions

We come back to the model theoretic context. We assume that our ambient theory T is NIP.

Recall from Theorem 3.1 (honest definitions) and the remark following it, that given $M \models T$, $A \subset M$ and $\phi(x; b) \in L(M)$, there is some $\psi(x; z) \in L$ such that for any finite $A_0 \subseteq \phi(A; b)$ we can find $d \in A$ with $A_0 \subseteq \psi(A; d) \subseteq \phi(A; b)$.

We now address the question of uniformity of ψ with respect to ϕ . First, compactness gives a weak uniformity statement.

Lemma 6.10. *Let $\phi(x; y) \in L$. For any formula $\psi(x; z)$ (where x is the same variable as in ϕ and z may vary), let an integer n_ψ be given. Then there are finitely many formulas $\psi_0, \dots, \psi_{n-1}$ such that:*

for $M \models T$, $A \subseteq M$, $b \in M$, there exists $j < n$, such that for any $A_0 \subseteq \phi(A; b)$ of size $\leq n_\psi$ there is some $d \in A$ with $A_0 \subseteq \psi_j(A; d) \subseteq \phi(A; b)$.

Proof. Consider the language $L' = L \cup \{\mathbf{P}(x); c_b\}$, where c_b is a new constant. Consider the L' -theory T' axiomatized by T along with the formulas Θ_ψ , for $\psi(x; z) \in L$, where

$$\begin{aligned} \Theta_\psi \quad \equiv \quad & (\exists x_0, \dots, x_{n_\psi-1} \in \mathbf{P}) \bigwedge_{i < n_\psi} \phi(x_i; c_b) \wedge \neg \exists d \in \mathbf{P} \\ & \bigwedge \psi(x_i; d) \wedge \forall x \in \mathbf{P} \psi(x; d) \rightarrow \phi(x; c_b). \end{aligned}$$

By Theorem 3.1 as recalled above, the theory T' is inconsistent. Therefore by compactness, there are finitely many formulas $\psi_0, \dots, \psi_{n-1}$ such that $T \cup \{\Theta_{\psi_0}, \dots, \Theta_{\psi_{n-1}}\}$ is inconsistent. Then we have what we want. \square

Theorem 6.11. *Let $\phi(x; y) \in L$. There exists $\psi(x; z) \in L$ such that for any $M \models T$, $A \subseteq M$ of size ≥ 2 , $b \in M$ and $A_0 \subseteq \phi(A; b)$ finite, there is $d \in A$ with*

$$A_0 \subseteq \psi(A; d) \subseteq \phi(A; b).$$

Proof. For any formula $\psi(x; z)$, let k_ψ be the VC-dimension of ψ and let $(n_\psi, N(\psi))$ be given by Corollary 6.8 for $k = k_\psi$. Apply the previous lemma to those n_ψ 's. It gives us formulas $\psi_0, \dots, \psi_{n-1}$. For $i < n$, let

$$\Psi_i(x; y_1, \dots, y_{N(\psi_i)}) = \bigvee_{j=1}^{N(\psi_i)} \psi_i(x; z_j).$$

Let now $M \models T$, $A \subseteq M$ and $b \in M$. By the previous lemma, there is some $i < n$ such that for any $A_0 \subseteq \phi(A; b)$ of size $\leq n_{\psi_i}$, we can find $d \in A$ with $A_0 \subseteq \psi(A; d) \subseteq \phi(A; b)$.

Let $A_1 \subseteq \phi(A; b)$ be finite. Let

$$\mathcal{S} = \{\psi_i(A_1; d) : \psi_i(A; d) \subseteq \phi(A; d)\} \subseteq \mathfrak{P}(A_1).$$

Then the set system \mathcal{S} has VC-dimension less than or equal to $\text{VC}(\psi_i)$. Furthermore, the assumptions on ψ_i imply that \mathcal{S} has the (n_{ψ_i}, n_{ψ_i}) -property. Therefore by Corollary 6.8, there are $S_1, \dots, S_{N(\psi_i)}$ which cover the whole of A_1 . Write $S_j = \psi_i(A_1; d_j)$. Then

$$A_1 \subseteq \Psi_i(A; d_1, \dots, d_{N(\psi_i)}) \subseteq \phi(A; b).$$

If $|A| \geq 2$, then by usual coding tricks, we can replace the finite set $\{\Psi_0, \dots, \Psi_{n-1}\}$ by a single formula, and the theorem follows. \square

The following corollary is usually referred to as “uniform definability of types over finite sets”, or UDTFS.

Corollary 6.12 (UDTFS). *Let $\phi(x; y) \in L$, then there is $\psi(x; z) \in L$ such that for any $b \in \mathcal{U}$ and $A \subset \mathcal{U}$ a finite set, there is $d \in A$ with*

$$\phi(A; b) = \psi(A; d).$$

Proof. Simply apply Theorem 6.11 with $A = A_0$. □

Note that this implies that the number of ϕ -types over A is bounded by $|A|^{|z|}$. One can see this as a model theoretic version of Sauer's lemma 6.3 which says that the number of ϕ -types is polynomial in the size of A . However we here have no a priori bound on $|z|$.

Remark 6.13. We have used the fact that the theory T is NIP to obtain an honest definition for $\phi(x; b)$ and again to apply Corollary 6.8 to the formulas $\psi(x; z)$. In particular, the proof does not go through assuming that only $\phi(x; y)$ is NIP. It is an open question whether or not UDTFS holds in that case.

Historical Remarks

VC-dimension was introduced in the context of statistical learning theory by Vapnik and Chervonenkis in [94] where Theorem 6.4 is proved. It has played a major role in that area. Sauer's lemma 6.3 is implicit in [94], and was rediscovered independently by at least two authors: Shelah [79] and Sauer [74]. Shelah and Sauer's proofs are by induction on n . The proof we give was found independently by Alon [7] and Frankl [31].

The (p, q) theorem was proved for families of convex subsets of the euclidian space by Alon and Kleitman [5] and adapted to families of finite VC-dimension by Matousek [62]. The udtfs property was conjectured by Laskowski and proved for o-minimal theories by him and Johnson in [48]. Then Guingona proved it for dp-minimal theories [33]. It was proved for all NIP theories in [22]. This problem is linked to a similar question in machine learning theory (existence of *compression schemes* for learnable families). See the discussion and references in [48].

Chapter 7

Measures

In this chapter we study *Keisler measures* which can be seen as a generalization of types which allow for truth values in the segment $[0, 1]$, or as ordinary probability measures on the compact space of types. The two points of view are useful and we will often switch from one to the other.

Many properties of types generalize naturally to measures. In particular, we will define generically stable measures. As we will see, those are ubiquitous in NIP theories, as opposed to generically stable *types*. One way to obtain such a measure is to take a σ -additive probability measure on a standard model (for example the Lebesgue measure on $[0, 1]$ seen as subset of the structure \mathbb{R}).

The main use of measures seems to be as invariant measures on group, but we will not explain that here and refer the reader to [47] and [42] for more information.

In this chapter, we again assume that T is NIP, although the basic definitions are of course valid in any theory.

7.1 Definitions and basic properties

If A is a set of parameters, and x a variable, we let $\mathcal{L}_x(A)$ denote the algebra of A -definable sets in the variable x . Equivalently, it is the boolean algebra of formulas with parameters in A and free variable x , quotiented by the equivalence relation $\phi(x) \sim \psi(x) \iff \mathcal{U} \models \phi(x) \leftrightarrow \psi(x)$.

Definition 7.1. Let $A \subset M$ be a set of parameters. A *Keisler measure* (or simply a *measure*) μ over A in the variable x is a finitely additive probability measure on $\mathcal{L}_x(A)$. In other words it is a function $\mu : \mathcal{L}_x(A) \rightarrow [0, 1]$ such that:

$$\begin{aligned}
\mu(x = x) &= 1; \\
\mu(\neg\phi(x)) &= 1 - \mu(\phi(x)); \\
\mu(\phi(x) \wedge \psi(x)) + \mu(\phi(x) \vee \psi(x)) &= \mu(\phi(x)) + \mu(\psi(x)).
\end{aligned}$$

We will sometimes write μ as μ_x or $\mu(x)$ to emphasize that μ is a measure on the variable x . If $A \subseteq B$, and $\mu \in \mathfrak{M}_x(B)$, we define the restriction of μ to A denoted $\mu|_A$ or $\mu \upharpoonright A$ as the restriction of μ to $\mathcal{L}_x(A)$. As for types, we say that μ is an extension to B of $\mu|_A$.

EXAMPLE 7.2. • If $p \in S_x(A)$ is a type over A , then p will be identified with the Keisler measure $\mu_p(x)$ over A defined by $\mu_p(\phi(x)) = 1$ when $p \vdash \phi(x)$. Thus a (complete) type is a special case of a Keisler measure.

We will often identify the type p with the associated measure, and write for example $p(\phi(x; b))$ which stands for $\mu_p(\phi(x; b))$.

- Given a_0, a_1, \dots in $[0, 1]$ such that $\sum a_i = 1$, and types p_0, p_1, \dots over A in the same variable x , we can define the average measure $\mu = \sum a_i p_i$.
- Take T to be the theory of real closed fields and let \mathbb{R} be the standard model. Let μ_0 be any Borel probability measure on \mathbb{R} . Then μ_0 induces a Keisler measure over \mathcal{U} in one variable x defined by $\mu(\phi(x)) = \mu_0(\phi(\mathbb{R}))$.
- Let $I = (a_i : i \in [0, 1])$ be an indiscernible sequence. Let λ_0 denote the usual Lebesgue measure on the interval $[0, 1]$. We can define the *average measure* $Av(I)$ as the measure μ defined by $\mu(\phi(x; b)) = \lambda_0(\{i \in [0, 1] : \models \phi(a_i; b)\})$. Note that NIP ensures that the set in question is Lebesgue measurable (it is in fact a finite union of intervals).
- Let $(M_n : n < \omega)$ be a sequence of finite structures. For each $n < \omega$, let μ_n denote the counting measure on M_n . Let \mathcal{D} be an ultrafilter on ω and consider the ultraproduct $M = \prod_{\mathcal{D}} M_n$. We define a measure μ on M in the following way. Let $\phi(x; b) \in L(M)$ be a formula. Let $(b_n : n < \omega)$ be a representative of b in $\prod_{n < \omega} M_n$. Then set $\mu(\phi(x; b)) = \lim_{\mathcal{D}} \mu_n(\phi(M_n; b_n))$.
- Here is an example outside of NIP that can serve as counter-example to most of the results in this section which require NIP. Let T be the theory of the random graph in the language $\{R\}$. Let $M \models T$. Define a Keisler measure μ on M by $\mu(\bigcap_{i < n} (x R a_i)^{\eta(i)}) = 2^{-n}$, where the a_i 's are pairwise distinct points of M and $\eta : n \rightarrow \{0, 1\}$ is any function.

Let $\mathfrak{M}_x(A)$ denote the set of Keisler measures over A . It is a closed subset of $[0, 1]^{\mathcal{L}_x(A)}$, equipped with the product topology. We equip $\mathfrak{M}_x(A)$ with the induced topology, making it a compact space. The identification of a type with the measure it defines gives an identification of $S_x(A)$ as a closed subspace of $\mathfrak{M}_x(A)$.

Borel measures

Let $\mu \in \mathfrak{M}_x(A)$ be a Keisler measure. It assigns a measure to every clopen set of the space $S_x(A)$. We show how to extend that measure to a σ -additive Borel probability measure. First, if $O \subseteq S_x(A)$ is open, we define $\mu(O) = \sup\{\mu(D) : D \subseteq O, D \text{ clopen}\}$. And similarly for closed sets. Note that if X is either closed or open, then we have

$$(Reg) \quad \sup\{\mu(F) : F \subseteq X, F \text{ closed}\} = \inf\{\mu(O) : X \subseteq O, O \text{ open}\}.$$

We show that the set of subsets $X \subseteq S_x(A)$ satisfying (Reg) is closed under complement and countable union. Complement is clear. For countable union: let $X = \bigcup_{i < \omega} X_i$ and fix $\epsilon > 0$. For each $i < \omega$, take $F_i \subseteq X_i \subseteq O_i$ with $\mu(O_i \setminus F_i) \leq \epsilon 2^{-i}$. Let $O = \bigcup_{i < \omega} O_i$. Then we can find some finite N such that $\mu(O \setminus \bigcup_{i < N} O_i) \leq \epsilon$. Let $F = \bigcup_{i < N} F_i$. Then we have $F \subseteq X \subseteq O$ and $\mu(O \setminus F) \leq 2\epsilon$.

It follows that every Borel subset of $S_x(A)$ satisfies (Reg) . We can therefore define μ on all such sets by $\mu(X) = \sup\{\mu(F) : F \subseteq X, F \text{ closed}\} = \inf\{\mu(O) : X \subseteq O, O \text{ open}\}$. It is easy to check that this defines a σ -additive measure on $S_x(A)$. Property (Reg) is referred to as *regularity* of the measure μ .

So, to a Keisler measure on A , we have associated a regular probability measure on $S_x(A)$. Conversely, if μ is a regular probability measure on $S_x(A)$, then it defines a Keisler measure by restriction to the clopens. Regularity ensures that μ is entirely determined by that restriction. We therefore obtain a bijection

$$\{\text{Keisler measures on } A\} \longleftrightarrow \{\text{Regular Borel probability measures on } S_x(A)\}.$$

We will often use that bijection implicitly, using the same notation for the Keisler measure and for the Borel measure on the type space. In particular, if μ is a Keisler measure on A , we may write $\mu(X)$, where X is a Borel subset of $S_x(A)$.

We define the *support* of μ as the set $S(\mu) \subseteq S_x(A)$ of types $p \in S_x(A)$ such that $p(\phi(x; b)) = 1 \implies \mu(\phi(x; b)) > 0$ for any $\phi(x; b) \in L(A)$. The support of μ is thus a closed set of $S_x(A)$. A type in the support of μ is called *weakly random* for μ .

If X is a Borel set of positive μ -measure, we define the localization of μ at X to be the measure $\mu_X(x)$ defined by $\mu_X(\phi(x)) = \mu(\phi(x) \cap X)/\mu(X)$.

Extending measures

The next lemma says that any partial measure in a suitable sense extends to a full Keisler measure.

Lemma 7.3. *Let $\Omega \subseteq \mathcal{L}_x(A)$ be a set of (equivalence classes of) formulas closed under intersection, union and complement and containing \top . Let μ_0 be a finitely additive measure on Ω with values in $[0, 1]$ such that $\mu_0(\top) = 1$. Then μ extends to a Keisler measure over A .*

Proof. By compactness in the space $[0, 1]^{\mathcal{L}_x(A)}$, it is enough to show that given $\psi_1(x), \dots, \psi_n(x)$ formulas in $\mathcal{L}_x(A)$, there is a function $f = \langle \psi_1, \dots, \psi_n \rangle \rightarrow [0, 1]$ finitely additive and compatible with μ_0 (where $\langle B \rangle$ denotes the boolean algebra generated by B). We may assume that ψ_1, \dots, ψ_n are the atoms of the boolean algebra B that they generate.

The elements of Ω in B form a sub-boolean algebra. Let ϕ_1, \dots, ϕ_m be its atoms. We have say:

$$\phi_1 = \psi_{i_1(1)} \vee \dots \vee \psi_{i_1(l_1)} \quad \phi_2 = \psi_{i_2(1)} \vee \dots \vee \psi_{i_2(l_2)} \quad \text{etc.}$$

Then any finitely additive f satisfying $f(\psi_{i_1(1)}) + \dots + f(\psi_{i_1(l_1)}) = \mu_0(\phi_1)$ etc. will do. \square

Lemma 7.4. *Let $\mu \in \mathfrak{M}_x(M)$ be a measure and let $\phi(x; b) \in L(\mathcal{U})$. Let $r_1 = \sup\{\mu(\psi(x)) : \psi(x) \in L(M), \models \psi(x) \rightarrow \phi(x; b)\}$ and $r_2 = \inf\{\mu(\psi(x)) : \psi(x) \in L(M), \models \phi(x; b) \rightarrow \psi(x)\}$.*

Then for any $r_1 \leq r \leq r_2$, there is an extension $\nu \in \mathfrak{M}(\mathcal{U})$ of μ such that $\nu(\phi(x; b)) = r$.

Proof. It is enough to find some ν_1, ν_2 such that $\nu_i(\phi(x; b)) = r_i$ for $i \in \{1, 2\}$, since we can then consider averages of ν_1 and ν_2 . We build ν_2 . Let $\Omega \subset \mathcal{L}_x(\mathcal{U})$ be the boolean algebra generated by $L(M)$ and the formula $\phi(x; b)$. By Lemma 7.3 it is enough to define ν_2 on Ω .

First assume that $r_2 = 0$. Then for any $\theta(x) \in L(M)$, set $\nu_2(\theta(x)) = \mu(\theta(x))$ and set $\nu_2(\phi(x; b)) = 0$. This extends uniquely to a finitely additive measure on Ω . Thus we are done. A similar argument works for $r_2 = 1$.

If $r_2 > 0$, then let μ' be the localization of μ on the closed set $C = \bigwedge \theta(x)$ where $\theta(x)$ ranges over formulas of $L(M)$ such that $\models \phi(x; b) \rightarrow \theta(x)$. Let μ'' be the localization of μ on the complementary open set. Then we have $\mu = r_2\mu' + (1 - r_2)\mu''$. By the previous paragraph, we can extend μ' and μ'' respectively to ν' and ν'' such that $\nu'(\phi(x; b)) = 1$ and $\nu''(\phi(x; b)) = 0$. Then set $\nu_2 = r_2\nu' + (1 - r_2)\nu''$.

The construction of ν_1 follows by taking $\neg\phi(x; b)$ instead of $\phi(x; b)$. \square

7.2 Boundedness properties

Lemma 7.5. *Let $\mu \in \mathfrak{M}_x(M)$ be a measure and $(b_i : i < \omega)$ an indiscernible sequence in M . Let $\phi(x; y)$ be a formula and $r > 0$ such that $\mu(\phi(x; b_i)) \geq r$ for all $i < \omega$. Then the partial type $\{\phi(x; b_i) : i < \omega\}$ is consistent.*

Proof. First we show that we can assume that the sequence (b_i) is μ -indiscernible, by which we mean that if $i_1 < \dots < i_n < \omega$ and $j_1 < \dots < j_n < \omega$, then $\mu(\phi(x; b_{i_1}) \wedge \dots \wedge \phi(x; b_{i_n})) = \mu(\phi(x; b_{j_1}), \dots, \phi(x; b_{j_n}))$.

Consider the set of conditions on a measure $\nu \in \mathfrak{M}_x(M)$:

- $\nu(\phi(x; b_i)) \geq r$, for all $i < \omega$;
- $|\nu(\phi(x; b_{i_1}) \wedge \dots \wedge \phi(x; b_{i_n})) - \nu(\phi(x; b_{j_1}), \dots, \phi(x; b_{j_n}))| \leq 1/N$, for every integer N and any $i_1 < \dots < i_n < \omega$ and $j_1 < \dots < j_n < \omega$.

Note that each one of these conditions defines a closed subset of $\mathfrak{M}_x(M)$. By Ramsey and the existence of μ , we see that any finite set of such conditions is consistent. Therefore by compactness in the space $\mathfrak{M}_x(M)$, the whole set is consistent, realized by some ν . Replacing μ by ν , we have what we want.

Assume for a contradiction that $\{\phi(x; b_i) : i < \omega\}$ is inconsistent. Then there is some N such that $\mu(\phi(x; b_0) \wedge \dots \wedge \phi(x; b_{N-1})) = 0$. Take a minimal such N and let $N' = N - 1$. For any integer m , let $\psi_m(x) = \phi(x; b_{mN'}) \wedge \dots \wedge \phi(x; b_{mN'+N'-1})$. Then by minimality of N and indiscernability of the sequence, there is some $t > 0$ such that we have $\mu(\psi_m(x)) = t$ for all m . Also by the property of N , we have $\mu(\psi_m(x) \wedge \psi_{m'}(x)) = 0$ for $m \neq m'$.

But then we have $\mu(\psi_0(x) \vee \dots \vee \psi_{m-1}(x)) = mt$, for all m . This contradicts the fact that the measure of the whole space is 1. \square

Lemma 7.6. *Let $\mu \in \mathfrak{M}_x(M)$. We cannot find a sequence $(b_i : i < \omega)$ of tuples of M , a formula $\phi(x; y)$ and $\epsilon > 0$ such that $\mu(\phi(x; b_i) \Delta \phi(x; b_j)) > \epsilon$ for all $i, j < \omega$, $i \neq j$.*

Proof. Assume otherwise. Then by Ramsey and compactness, we may assume that the sequence $(b_i : i < \omega)$ is indiscernible. By NIP, the partial type $\{\phi(x; b_{2k}) \Delta \phi(x; b_{2k+1}) : k < \omega\}$ is inconsistent. This contradicts the previous lemma. \square

Let $\mu \in \mathfrak{M}_x(M)$. If $X(x)$ and $Y(x)$ are two M -definable sets, set $X \sim_\mu Y$ if $\mu(X \Delta Y) = 0$. Then the previous lemma implies that the set of \sim_μ -equivalence classes has bounded cardinality. In particular, the support $S(\mu)$ of μ has cardinality bounded independently of M .

7.3 Smooth measures

We now define and investigate the notion of a smooth measure, which can be considered as an analog of realized types.

Definition 7.7. Let $\mu \in \mathfrak{M}_x(M)$. We say that μ is *smooth* if for every $N \supseteq M$, μ has a unique extension to an element of $\mathfrak{M}_x(N)$.

More generally, if $\mu \in \mathfrak{M}_x(N)$ and $M \subseteq N$, we say that μ is smooth over M if $\mu|_M$ is smooth.

Lemma 7.8. *Let $\mu \in \mathfrak{M}_x(M)$ be a smooth measure. Let $\phi(x, y) \in L$ and $\epsilon > 0$. Then there are formulas $\theta_i^1(x), \theta_i^2(x)$ for $i = 1, \dots, n$ and $\psi_i(y)$ for $i = 1, \dots, n$, all over M such that:*

- (i) *the formulas $\psi_i(y)$ partition y -space;*
- (ii) *for all i , if $\models \psi_i(b)$, then $\models \theta_i^1(x) \rightarrow \phi(x, b) \rightarrow \theta_i^2(x)$;*
- (iii) *for each i , $\mu(\theta_i^2(x)) - \mu(\theta_i^1(x)) < \epsilon$.*

Proof. Let $b \in \mathcal{U}$. Then by smoothness of μ and Lemma 7.4, there are formulas $\theta_i^1(x), \theta_i^2(x)$ over M , such that

(*) $\models \theta_i^1(x) \rightarrow \phi(x, b) \rightarrow \theta_i^2(x)$, and $\mu(\theta_i^2(x)) - \mu(\theta_i^1(x)) < \epsilon$.

By compactness, there are finitely many such pairs, say, $(\theta_i^1(x), \theta_i^2(x))$ such that for every b one of these pairs satisfies (*). It is then easy to find the $\psi_i(y)$. \square

Note that conversely, the existence of the formulas $\theta_i^t(x)$ and $\psi_i(y)$ with the above properties implies that the measure μ is smooth.

Proposition 7.9. *Let $\mu \in \mathfrak{M}_x(M)$ be any measure. Then there is $M \prec N$ and an extension $\mu' \in \mathfrak{M}_x(N)$ of μ which is smooth.*

Proof. Assume not. We build an increasing sequence $(M_\alpha : \alpha < |T|^+)$ of models, with $M_0 = M$ and associated increasing sequence of measures $\mu_\alpha \in \mathfrak{M}_x(M_\alpha)$. At limit stages, take the union. Assume M_α, μ_α are defined. By hypothesis, μ_α is not smooth, so we can find some $\phi_\alpha(x; b_\alpha) \in L(\mathcal{U})$ and μ^1, μ^2 two extensions of μ_α such that $\mu^2(\phi_\alpha(x; b_\alpha)) - \mu^1(\phi_\alpha(x; b_\alpha)) = 4\epsilon_\alpha > 0$. Then take $M_{\alpha+1}$ an extension of M_α containing b_α and set $\mu_{\alpha+1} = 1/2(\mu^1 + \mu^2)$. Note that $\mu_{\alpha+1}$ has the following property: for any $\theta(x) \in L(M_\alpha)$, $\mu_{\alpha+1}(\theta(x) \Delta \phi(x; b_\alpha)) \geq \epsilon_\alpha$.

Having done the construction, we may assume that $\phi_\alpha = \phi$, and $\epsilon_\alpha = \epsilon > 0$ are both constant. Let μ' be the measure $\bigcup_{\alpha < |T|^+} \mu_\alpha$. Then we have

$$\mu'(\phi(x; b_\alpha) \Delta \phi(x; b_\beta)) \geq \epsilon$$

for any $\alpha < \beta$. This contradicts Lemma 7.6. \square

We now show that smooth measures in NIP theories can be approximated by averages of points.

If μ_0, \dots, μ_{n-1} are measures and $\phi(x; b)$ is a formula, then the notation $\text{Av}(\mu_0, \dots, \mu_{n-1}; \phi(x; b))$ stands for $\frac{1}{n} \sum_{k < n} \mu_k(\phi(x; b))$.

Similarly, if a_0, \dots, a_{n-1} are tuples, then $\text{Av}(a_0, \dots, a_{n-1}; \phi(x; b))$ stands for $\frac{1}{n} |\{i : \models \phi(a_i; b)\}|$. We extend this notation naturally to the case where $\phi(x; b)$ is replaced by an arbitrary Borel subset of some type space.

Proposition 7.10. *Let μ_x be a global measure, smooth over M . Let X be a Borel subset of $S_x(M)$, and $\phi(x; y)$ a formula. Fix $\epsilon > 0$. Then there are $a_0, \dots, a_{n-1} \in \mathcal{U}$ such that for any $b \in \mathcal{U}$,*

$$|\mu(X \cap \phi(x; b)) - \text{Av}(a_0, \dots, a_{n-1}; X \cap \phi(x; b))| \leq \epsilon.$$

Proof. Fix formulas $\psi_i(y)$, $\theta_i^0(x)$ and $\theta_i^1(x)$, $i < m$ as given by Lemma 7.8. Consider μ as a probability measure on the space $S_x(M)$. Then we can find types $(p_i : i < n)$ in $S_x(M)$ such that if $\lambda = \frac{1}{n} \sum_{i < n} p_i$ then $\lambda(\theta_i^0(x) \cap X)$, $\lambda(\theta_i^1(x) \cap X)$ are all within ϵ of the corresponding $\mu(\theta_i^0(x) \cap X)$, $\mu(\theta_i^1(x) \cap X)$.

Now take points $(a_i : i < n)$ in \mathcal{U} such that $a_i \models p_i$. Set $\lambda' = \frac{1}{n} \sum_{i < n} \text{tp}(a_i/\mathcal{U})$. Let $b \in \mathcal{U}$ and let $i < n$ be such that $\models \psi_i(b)$. Then we have $\models \theta_i^0(x) \rightarrow \phi(x; b) \rightarrow \theta_i^1(x)$ and $\mu(\theta_i^1(x)) - \mu(\theta_i^0(x)) \leq \epsilon$. This implies that $\lambda(\theta_i^1(x) \cap X) - \lambda(\theta_i^0(x) \cap X) \leq 3\epsilon$ and therefore $\lambda'(\phi(x; b) \cap X)$ is within 3ϵ of $\lambda(\phi(x; b) \cap X)$ and thus within 4ϵ of $\mu(\phi(x; b) \cap X)$. \square

We conclude that we can approximate any measure by averages of types.

Proposition 7.11. *Let $\mu \in \mathfrak{M}_x(A)$ be any Keisler measure, and let $\phi(x; y) \in L$ be a formula and fix $X_1, \dots, X_m \subseteq S_x(A)$ Borel subsets. Let $\epsilon > 0$. Then there are types $p_0, \dots, p_{n-1} \in S_x(A)$ such that, for every $b \in A$ and every $k \leq m$:*

$$|\mu(\phi(x; b) \cap X_k) - \text{Av}(p_0, \dots, p_{n-1}; \phi(x; b) \cap X_k)| \leq \epsilon.$$

Furthermore, one can impose that $p_k \in S(\mu)$ for all k .

Proof. Let ν_x be a smooth extension of μ over a model $N \supset A$. Let $X \subseteq S_x(A)$ be the support of μ . The previous proposition works equally well with finitely many Borel subsets instead of one and gives points $a_0, \dots, a_{n-1} \in N$ such that $|\nu(\phi(x; b) \cap X_k) - \text{Av}(a_0, \dots, a_{n-1}; \phi(x; b) \cap X_k)| \leq \epsilon$ for all $b \in N$ and $k \leq m$. Then set $p_i = \text{tp}(a_i/A)$.

The “furthermore” part follows by taking $S(\mu)$ to be one of the X_k ’s. \square

EXERCISE 7.12. *Give another proof of Proposition 7.11 using the VC-theorem instead of smooth measures. Deduce that the number n of types can be chosen so as to depend only on $\text{VC-dim}(\phi(x; y))$ and ϵ (see [47, Lemma 4.8]).*

7.4 Invariant measures

We now extend a number of definitions from types to measures.

Definition 7.13. Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be a measure, and let $A \subset \mathcal{U}$. We say that μ is A -invariant if for every $b \equiv_A b'$, and $\phi(x; y) \in L$, we have $\mu(\phi(x; b)) = \mu(\phi(x; b'))$.

Definition 7.14. Let $A \subseteq B$ and $\mu \in \mathfrak{M}_x(B)$. Then μ *does not fork* (resp. divides) over A if $\mu(\phi(x; b)) = 0$ for every formula $\phi(x; b) \in L(B)$ which forks (resp. divides) over A .

Proposition 7.15. *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ and $A \subset \mathcal{U}$. Then μ does not fork over A if and only if it is $\text{Lstp}(A)$ -invariant.*

Proof. Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be $\text{Lstp}(A)$ -invariant. As it is a global measure, it is enough to show that it does not divide over A . Let $\phi(x; b) \in L(M)$ be such that $\mu(\phi(x; b)) > 0$. Let $(b_i : i < \omega)$ be an A -indiscernible sequence with $b_0 = b$. Then by assumption $\mu(\phi(x; b_i)) = \mu(\phi(x; b))$ for all $i < \omega$. By Lemma 7.5, the

partial type $\{\phi(x; b_i) : i < \omega\}$ is consistent. It follows that $\phi(x; b)$ does not divide over A .

Conversely, assume that μ does not fork over A . If $\text{Lstp}(b/A) = \text{Lstp}(b'/A)$, then the formula $\phi(x; b) \triangle \phi(x; b')$ forks over A . Thus $\mu(\phi(x; b) \triangle \phi(x; b')) = 0$. \square

Note in particular that, if $A = M$ is a model, and μ does not fork over M , then we have $\mu(\phi(x; b) \triangle \phi(x; b')) = 0$ whenever $b \equiv_M b'$.

Definition 7.16. Let $M \models T$ and let $\mu_x \in \mathfrak{M}(\mathcal{U})$:

- μ is *finitely satisfiable* in M if for every $\phi(x; b) \in L(\mathcal{U})$ with $\mu(\phi(x; b)) > 0$, there is $a \in M$ such that $\mathcal{U} \models \phi(a; b)$.
- μ is *definable* over M if it is M -invariant and for every $\phi(x; y) \in L$, and $r \in [0, 1]$, the set $\{p \in S_y(M) : \mu(\phi(x; b)) < r \text{ for any } b \in \mathcal{U}, b \models p\}$ is an open subset of $S_y(M)$.
- μ is *Borel-definable* over M if it is M -invariant and the above set is a Borel set of $S_y(M)$.

Note that if μ is finitely satisfiable in M then it is M -invariant. Also μ is definable over M if and only if it is M -invariant and for any open (resp. closed) subset X of $[0, 1]$, the set $\{p \in S_y(M) : \mu(\phi(x; b)) \in X \text{ for any } b \in \mathcal{U}, b \models p\}$ is an open (resp. closed) subset of $[0, 1]$.

A measure μ is M -invariant (resp. finitely satisfiable in M), if and only if all types in $S(\mu)$ are M -invariant (resp. finitely satisfiable in M).

Lemma 7.17. Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ and $M \prec \mathcal{U}$.

- (i) If μ is smooth over M , then μ is finitely satisfiable and definable in M .
- (ii) If μ is M -invariant and smooth, then it is smooth over M .

Proof. (i): Let $\phi(x; b)$ be such that $\mu(\phi(x; b)) > 0$. Then by lemma 7.8, taking ϵ small enough, there is a formula $\theta^0(x) \in L(M)$ such that $\models \theta^0(x) \rightarrow \phi(x; b)$ and $\mu(\theta^0(x)) > 0$. In particular, $\theta^0(x)$ is consistent, and therefore has a point in M . It follows that also $\phi(x; b)$ has a point in M .

If $\phi(x; y)$ and $\epsilon > 0$ are fixed, then again apply Lemma 7.8 to obtain formulas $\psi_i^\epsilon(y) \in L(M)$, $i < n$ giving a partition of y -space. Let $r \in [0, 1]$ be such that $\mu(\phi(x; b)) < r$. Let $\epsilon > 0$ such that $\mu(\phi(x; b)) < r - 2\epsilon$. Take $\psi_i^\epsilon(y)$ satisfied by b . Then for any b' , $\models \psi_i^\epsilon(y)$ implies that $\mu(\phi(x; b')) \leq r - \epsilon < r$. This proves definability of μ .

(ii): Assume that μ is M -invariant and smooth. First note that, as in Lemma 2.21, μ is definable over M and finitely satisfiable in M . Fix a formula $\phi(x; y)$ and $\epsilon > 0$. Lemma 7.8 gives us formulas $\psi_i(y; d)$, $\theta_i^0(x; d)$ and $\theta_i^1(x; d)$, $i < n$, where we have made the parameters explicit. By definability of μ , there is an M -formula $\zeta(z)$ satisfied by d and such that for any $d' \in \zeta(\mathcal{U})$, the formulas $\psi_i(y; d')$, $\theta_i^0(x; d')$ and $\theta_i^1(x; d')$, $i < n$ also satisfy the conclusion of the lemma. Then we can find such a d' in M . It follows that μ is smooth over M . \square

Lemma 7.18. *Let $p \in S(\mathcal{U})$ be some M -invariant type, then p is Borel-definable over M .*

Proof. Let $\phi(x; y)$ be a formula. We need to see that the set $\{b \in \mathcal{U} : p \vdash \phi(x; b)\}$ is Borel over M . Let $b \in \mathcal{U}$. Take a maximal N such that there is a sequence $(a_i : i < N) \models p^{(N)}|_M$ with

$$(*) \quad \neg(\phi(a_i; b) \leftrightarrow \phi(a_{i+1}; b)) \text{ for } i < N - 1.$$

By the discussion on eventual types in Section 2.2.2, we know that $\phi(a_{N-1}; b)$ holds if and only if $p \vdash \phi(x; b)$.

Let $A_N(b)$ say that there are $(a_i : i < N)$ such that $(*)$ holds and $\models \phi(a_{N-1}; b)$. Similarly we say that $B_N(b)$ holds if and there is $(a_i : i < N)$ such that $(*)$ holds and $\models \neg\phi(a_{N-1}; b)$. Then both $A_N(x)$ and $B_N(x)$ are type-definable sets. We have $p \vdash \phi(x; b)$ if and only if there is some integer $N \leq \text{alt}(\phi) + 1$ such that $A_N(x) \wedge \neg B_{N+1}(x)$ holds.

This shows that $\{b \in \mathcal{U} : p \vdash \phi(x; b)\}$ is a finite boolean combination of closed sets, definable over M . \square

Proposition 7.19. *Let $\mu \in \mathfrak{M}(\mathcal{U})$ be M -invariant, then μ is Borel-definable over M .*

Proof. This follows easily from the previous lemma and Proposition 7.11 which says that a measure can be approximated by types in its support. \square

We now have all we need to define the product of an invariant measure over another, similarly as we did for types. So let $\mu_x \in \mathfrak{M}(\mathcal{U})$ be M -invariant and let $\lambda_y \in \mathfrak{M}(\mathcal{U})$ be any measure. We define $\omega_{xy} = \mu_x \otimes \lambda_y$ as a global measure in two variables by the formula:

$$\omega_{xy}(\phi(x, y; b)) = \int_{p \in S_y(N)} f(p) d\lambda_x|_N,$$

where $f : p \mapsto \mu(\phi(x, d; b))$ for some (any) $d \models p$ and N is a small model containing Mb . Note that f is a measurable function by Borel-definability of μ . We need to see that this definition does not depend on the choice of N . So let $N_1 \subseteq N_2$ be two models containing Mb and let f_1, f_2 be the associated functions. Let $\pi : S(N_2) \rightarrow S(N_1)$ be the canonical restriction map. Then $f_2 = \pi^{-1}(f_1)$. For any clopen set $B \subseteq N_1$, $\lambda|_{N_1}(B) = \lambda|_{N_2}(\pi^{-1}(B))$, hence this holds also when B is Borel and then by approximating f_1 by step functions, we deduce that the integral of f_1 and f_2 are the same.

Note that if $\mu = p$ and $\lambda = q$ are types, then we recover the usual product $p_x \otimes q_y$.

If λ is itself M -invariant, then $\mu_x \otimes \lambda_y$ is M -invariant. Also if both μ and λ are finitely satisfiable in M (resp. M -definable), then $\mu_x \otimes \lambda_y$ is finitely satisfiable in M (resp. M -definable).

We now argue that the product of measures is associative. Let μ_x, η_y and λ_z be three global measures, and assume that μ and η are both M -invariant. Let $\phi(x, y, z) \in L(M)$, let $a = ((\mu_x \otimes \eta_y) \otimes \lambda_z)(\phi(x, y, z))$ and $b = (\mu_x \otimes (\eta_y \otimes \lambda_z))(\phi(x, y, z))$. We want to show that $a = b$.

First, assume that $\mu = p$ is a type. Let $d\phi(y, z)$ denote the Borel subset of $S_{y,z}(M)$ defined by $r \in d\phi \iff p \vdash \phi(x, b, c)$ where $(b, c) \models r$. Then $b = \eta_y \otimes \lambda_z(d\phi(y, z)) = \int_{q \in S_{y,z}(M)} f(p) d\lambda$, where $f : q \rightarrow \eta(d\phi(y, c))$, for some $c \models q$. But this is exactly a .

Now if μ is an arbitrary invariant measure, then given a formula $\phi(x, y, z)$ and ϵ , we can find a measure $\tilde{\mu}$, which is an average of types such that $|\mu(\phi(x, b, c)) - \tilde{\mu}(\phi(x, b, c))| \leq \epsilon$ for all b, c . Let \tilde{a}, \tilde{b} be the corresponding a and b with μ replaced by $\tilde{\mu}$. Then \tilde{a} and \tilde{b} are respectively at distance ϵ from a and b . By the previous paragraph, $\tilde{a} = \tilde{b}$, and we conclude that $a = b$.

If μ is an M -invariant global measure, then we can define by induction $\mu^{(1)}(x) = \mu(x)$ and $\mu^{(n+1)}(x_0, \dots, x_n) = \mu(x_0) \otimes \mu^{(n)}(x_1, \dots, x_n)$. And of course $\mu^{(\omega)}(x_0, x_1, \dots)$ is the union of $\mu^{(n)}(x_0, \dots, x_{n-1})$ for $n < \omega$.

Similarly as for types, we prove that a finitely satisfiable measure and a definable one commute.

Lemma 7.20. *Let μ_x, λ_y be two global M -invariant measures. Assume that $\mu_x \otimes p_y = p_y \otimes \mu_x$ for any $p \in S_y(\mathcal{U})$ in the support of λ . Then μ and λ commute.*

Proof. Let $\phi(x, y)$ be a formula, and write $(\mu_x \otimes \lambda_y)(\phi(x, y)) = \int_{p \in S_y(M)} f(p) d\lambda_x$ as in the definition. Fix $\epsilon > 0$. Approximate that integral up to ϵ by a finite sum

$\sum_{i=1,\dots,n} \lambda(X_i) c_i$, where $X_i = \{q \in S_y(M) : \mu(\phi(x, b)) \in [r_i, t_i], \text{ for some } b \models q\}$ for some $r_i, t_i \in [0, 1]$. By Proposition 7.11, we can find types q_1, \dots, q_m weakly random for λ , such that if $\tilde{\lambda}$ denotes the average $\frac{1}{m} \sum q_j$, then:

- $\tilde{\lambda}(X_i)$ is within ϵ of $\lambda(X_i)$, for all $i \leq n$;
- $\tilde{\lambda}(\phi(a, y))$ is within ϵ of $\lambda(\phi(a, y))$ for all $a \in \mathcal{U}$.

The first condition ensures that $\mu_x \otimes \tilde{\lambda}_y(\phi(x, y))$ is within ϵ of $\mu_x \otimes \lambda_y(\phi(x, y))$ and the second one ensures that $\tilde{\lambda}_y \otimes \mu_x(\phi(x, y))$ is within ϵ of $\lambda_y \otimes \mu_x(\phi(x, y))$. As μ commutes with $\tilde{\lambda}$, the result follows. \square

Proposition 7.21. *Let $\mu_x \in \mathfrak{M}(\mathcal{U})$ be definable and $\lambda_y \in \mathfrak{M}(\mathcal{U})$ be finitely satisfiable, then $\mu_x \otimes \lambda_y = \lambda_y \otimes \mu_x$.*

Proof. By the previous lemma, we may assume that $\lambda_y = q_y$ is a type. Let M such that both μ and q are M -invariant. Assume for a contradiction that there is $r \in [0, 1]$ and $\epsilon > 0$ such that $\mu_x \otimes q_y(\phi(x; y)) < r - \epsilon$ whereas $q_y \otimes \mu_x(\phi(x; y)) > r + \epsilon$. By definability of μ , there is an M -formula $\psi(y)$ such that $q \vdash \psi(y)$ and for all $b \in \psi(\mathcal{U})$, we have $\mu(\phi(x; b)) < r$. Pick $p_1, \dots, p_n \in S(M)$ such that for every $b \in M$, $\text{Av}(p_1, \dots, p_n; \phi(x; b))$ is strictly within ϵ of $\mu(\phi(x; b))$. Realize each p_i by $a_i \in \mathcal{U}$. By finite satisfiability of q , find some $b \in \psi(M)$ such that $\models \phi(a_i; b) \iff q \vdash \phi(x; b)$ for all i .

By choice of p_i 's and $\psi(y)$, we have $\text{Av}(p_1, \dots, p_n; \phi(x; b)) > r$. But that quantity is equal to $\frac{1}{n} \sum_i q(\phi(a_i; y))$, by choice of b . And therefore must be $> r$. Contradiction. \square

7.5 Generically stable measures

Similarly as we did for types, we define a *generically stable measure* as being a global measure which is both definable and finitely satisfiable (in some small model M). We will prove an analog of Theorem 2.33. First we define a new notion.

Definition 7.22. Let μ_x be a global M -invariant measure. We say that μ is *fm* (frequency interpretation measure) if for any formula $\phi(x; y) \in L$ and $\epsilon > 0$, there is a family $(\theta_n(x_1, \dots, x_n) : n < \omega)$ of formulas in $L(M)$ such that:

- $\lim \mu^{(n)}(\theta_n(x_1, \dots, x_n)) = 1$;
- for any $(a_1, \dots, a_n) \in \theta_n(\mathcal{U})$, and any $b \in \mathcal{U}$, $\text{Av}(a_1, \dots, a_n; \phi(x; b))$ is within ϵ of $\mu(\phi(x; b))$.

We will see later that a measure is fin if and only if it is generically stable. Left to right is easy, but the converse requires more work.

The main result needed is an adaptation of the VC-theorem from measure theory to the context of Keisler measures.

VC-type results

We call a measure $\mu(x_1, \dots, x_n)$ over A *symmetric* if for any permutation σ of $\{1, \dots, n\}$ and any formula $\phi(x_1, \dots, x_n) \in L(A)$, we have $\mu(\phi(x_1, \dots, x_n)) = \mu(\phi(x_{\sigma 1}, \dots, x_{\sigma n}))$. We extend naturally this definition to measures in infinitely many variables.

In the following, we fix some definable set \mathcal{X} over which variables x, x_1, \dots range. Let also $\phi(x; y)$ be a formula. Let n be an integer, and write $\bar{x} = (x_1, \dots, x_n)$ and $\bar{x}' = (x'_1, \dots, x'_n)$. We define

$$\delta'_n(\bar{x}, \bar{x}'; b) = |\text{Av}(x_1, \dots, x_n; \phi(x; b)) - \text{Av}(x'_1, \dots, x'_n; \phi(x; b))|$$

and

$$\delta_n(\bar{x}, \bar{x}') = \sup_{b \in \mathcal{U}} \delta'_n(\bar{x}, \bar{x}'; b).$$

Note that for any $\epsilon > 0$ the statement $\delta'_n(\bar{x}, \bar{x}'; b) \leq \epsilon$ is a definable property of $(\bar{x}, \bar{x}'; b)$. Similarly, the statement $\delta_n(\bar{x}, \bar{x}') \leq \epsilon$ is a definable property of (\bar{x}, \bar{x}') .

Lemma 7.23. *Let $\phi(x; y)$ be a formula, and fix $\epsilon > 0$. For each $n < \omega$, let $\mu_n(x_1, \dots, x_n, x'_1, \dots, x'_n)$ be a symmetric measure on \mathcal{X}^{2n} over \emptyset . Let E_n denote the μ_n -expectancy of δ_n . Then $E_n \rightarrow 1$.*

Proof. The proof is essentially the same as that of the claim inside the proof of Theorem 6.4. As done there, we set $Y = (\mathbb{Z}_2)^n$ and consider its action on \mathcal{X}^{2n} by flipping variables. We fix $\epsilon > 0$ and define as in 6.4 the set $A_* \subseteq Y$ of bad permutations, which has measure at most $2\pi_\phi(2n) \exp(-n\epsilon^2/8)$, where π_ϕ is the shatter function of the family of sets $\{\phi(\mathcal{U}; b) : b \in \mathcal{U}\}$. For $\sigma \notin A_*$, we have $\delta_n(\sigma(\bar{x}, \bar{x}')) \leq \epsilon/2$. For $\sigma \in A_*$ we have at any rate $\delta_n(\sigma(\bar{x}, \bar{x}')) \leq 1$. Thus for n big enough, for any \bar{x}, \bar{x}' ,

$$2^{-n} \sum_{\sigma \in Y} \delta_n(\sigma(\bar{x}, \bar{x}')) \leq \epsilon.$$

By the symmetry of μ_n , E_n equals the μ_n -expectation of $\sup_b \delta_n(\sigma(\bar{x}, \bar{x}'))$ for any $\sigma \in Y$, hence it is also equal to the expectation of the average. Therefore for n big enough, $E_n \leq \epsilon$. \square

Corollary 7.24. *Let $\mu(x_1, \dots)$ be a Keisler measure in ω variables over some set A . Assume that $\mu|_{\emptyset}$ is symmetric. Let $\phi(x; y) \in L$ and fix $\epsilon > 0$. Then there is n such that $\mu(\exists y(\delta'_n(x_1, \dots, x_n, x_{n+1}, \dots, x_{2n}; y) \geq \epsilon)) < \epsilon$.*

Proof. Apply Lemma 7.23 letting $\mu_n(x_1, \dots, x_{2n})$ be the restriction of μ to the variables (x_1, \dots, x_{2n}) . \square

Proposition 7.25. *Let μ be a global M -invariant measure, and assume that $\mu^{(\omega)}(x_1, \dots)|_{\emptyset}$ is symmetric. Then for any formula $\phi(x; y)$ and $\epsilon > 0$, there are $a_1, \dots, a_n \in \mathcal{U}$ such that for all $b \in \mathcal{U}$, $\text{Av}(a_1, \dots, a_n; \phi(x; b))$ is within ϵ of $\mu(\phi(x; b))$.*

Proof. Fix a formula $\phi(x; y)$ and $\epsilon > 0$. Write $\bar{a} = (a_1, \dots, a_n)$ and $\bar{a}' = (a'_1, \dots, a'_n)$. Define $\theta_n(\bar{a}, \bar{a}')$ to say that for all b , $\text{Av}(a_1, \dots, a_n; \phi(x; b))$ is within ϵ of $\text{Av}(a'_1, \dots, a'_n; \phi(x; b))$. For n big enough, we have the inequality $\mu^{(2n)}(\theta(x_1, \dots, x_{2n})) > 3/4$. In particular, there is $\bar{a} \in \mathcal{U}$, such that we have $\mu^{(n)}(\theta(\bar{a}, x_1, \dots, x_n)) > 3/4$.

On the other hand, let $\zeta_n(x_1, \dots, x_n)$ say that $\text{Av}(x_1, \dots, x_n; \phi(x; b))$ is within ϵ of $\mu(\phi(x; b))$. Then, since the variance of $\mu(\phi(x; b))$ is at most 1, Chebychev's inequality (C.2) gives $\mu^{(n)}(\zeta_n(x_1, \dots, x_n)) \leq 1/n\epsilon^2$. Taking n large enough, we have $\mu^{(n)}(\zeta_n(x_1, \dots, x_n)) \leq 1/2$.

It follow that $\mu^{(n)}(\theta(\bar{a}, x_1, \dots, x_n) \wedge \zeta_n(x_1, \dots, x_n)) > 0$. In particular $\text{Av}(a_1, \dots, a_n; \phi(x; b))$ is within 2ϵ of $\mu(\phi(x; b))$, as required. \square

We can do the same incorporating a Borel set X .

Proposition 7.26. *Let μ be a global M -invariant measure, and assume that $\mu^{(\omega)}(x_1, \dots)|_{\emptyset}$ is symmetric. Then for any Borel set X , formula $\phi(x; y)$ and $\epsilon > 0$, there are $a_1, \dots, a_n \in \mathcal{U}$ such that for all $b \in \mathcal{U}$, $\text{Av}(a_1, \dots, a_n; \phi(x; b) \cap X)$ is within ϵ of $\mu(\phi(x; b) \cap X)$.*

Proof. We need to go through the proof of Lemma 7.23 and of Proposition 7.25 replacing everywhere $\phi(x; b)$ by $\phi(x; b) \cap X$. This poses no difficulty. \square

Remark 7.27. The statements above imply that if μ is A -invariant and $\mu^{(\omega)}$ is symmetric, then μ is equal to the average of any *realization* of $\mu^{(\omega)}|_A$. More precisely, let $\lambda(x_1, \dots)$ be a global measure which extends $\mu^{(\omega)}|_A$. Let $\phi(x; b) \in L(\mathcal{U})$ be a formula and let $r = \mu(\phi(x; b))$. Fix $\epsilon > 0$, then for all but finitely many values of n , we have $\lambda(\phi(x_n; b)) \in (r - \epsilon, r + \epsilon)$.

This implies in particular that μ is determined by $\mu^{(\omega)}|_A$. (This is also true for arbitrary invariant measures, but we do not have the tools to prove it).

Proof. Assume that we have $\lambda(\phi(x_n; b)) - r > \epsilon$ for infinitely many values of n . Without loss, this is true for all n . Now set $\omega = \mu^{(\omega)}(x'_1, x'_2, \dots) \otimes \lambda(x_1, x_2, \dots)$. Then $\omega|_{\emptyset}$ is symmetric. By Corollary 7.24, for n large enough, we have $\omega(\delta'_n(x_1, \dots, x_n, x'_1, \dots, x'_n; b) < \epsilon/4) > 1/2$.

Let $\theta_n(x_1, \dots, x_n)$ say $\text{Av}(x_1, \dots, x_n; \phi(x; b)) - r > \epsilon/2$. Then by the weak law numbers, we have $\lambda(\theta_n(x_1, \dots, x_n)) \rightarrow 1$. Letting $\theta'_n(x'_1, \dots, x'_n)$ say $|\text{Av}(x'_1, \dots, x'_n; \phi(x; b)) - r| < \epsilon/4$, then we also have $\mu^{(\omega)}(\theta'_n(x'_1, \dots, x'_n)) \rightarrow 1$. It follows that $\omega(\theta_n(x_1, \dots, x_n) \wedge \theta'_n(x'_1, \dots, x'_n)) \rightarrow 1$. But this contradicts what we established in the previous paragraph. \square

Properties of generically stable measures

Theorem 7.28. *Let μ be a global M -invariant measure. Then the following are equivalent:*

- (i) μ is generically stable;
- (ii) μ is fm;
- (iii) $\mu^{(\omega)}(x_1, \dots)|_M$ is symmetric;
- (iv) we have $\mu_x \otimes \mu_y = \mu_y \otimes \mu_x$.

Proof. (i) \Rightarrow (iv): Follows from Proposition 7.21.

(iv) \Rightarrow (iii): Clear by associativity of \otimes .

(iii) \Rightarrow (ii): Follows from Proposition 7.25 (and its proof).

(ii) \Rightarrow (i): Easy. \square

Proposition 7.29. *Let μ_x be a generically stable measure, then for any invariant measure λ_y , we have $\mu_x \otimes \lambda_y = \lambda_y \otimes \mu_x$.*

Proof. By Lemma 7.20, we may assume that $\lambda_y = q_y$ is an invariant type. Let M be such that both μ and q are invariant over M and let $N \supset M$ be $|M|^+$ -saturated. Take some $b \models q|_N$ in \mathcal{U} . Let $\phi(x, y) \in L(M)$ be a formula. Let $X \subset S_x(M)$ be the set of types p such that $p \vdash \phi(a, y)$ for some (any) $a \models p$. Then, by definition, we have $\mu_x \otimes q_y(\phi(x, y)) = \mu(\phi(x, b))$ and $q_y \otimes \mu_x(\phi(x, y)) = \mu(X)$.

Fix $\epsilon > 0$. By Proposition 7.26, let $a_1, \dots, a_n \in N$ be such that for all $b' \in \mathcal{U}$, $\text{Av}(a_1, \dots, a_n; \phi(x; b'))$ is within ϵ of $\mu(\phi(x; b'))$ and $\text{Av}(a_1, \dots, a_n; X)$ is within ϵ of $\mu(X)$. Then $\mu_x \otimes q_y(\phi(x, y))$ is within ϵ of $\text{Av}(a_1, \dots, a_n; \phi(x; b))$, which by definition of X is within ϵ of $q_y \otimes \mu_x(\phi(x, y))$.

As this holds for all $\epsilon > 0$, the result follows. \square

Proposition 7.30. *Let μ_x be generically stable and A -invariant. Then μ is the unique A -invariant extension of $\mu|_A$.*

Proof. Assume that λ_x is a global A -invariant extension of μ . As in the proof of Proposition 2.39, we show that $\mu^{(\omega)}|_A = \lambda^{(\omega)}|_A$. It follows that λ is also generically stable, and by the remark after Proposition 7.26, $\mu = \lambda$. \square

EXAMPLE 7.31. Recall the examples of measures given in Example 7.2. We go through them again, and see that most of them are generically stable.

- If $p \in S_x(\mathcal{U})$ is a global type, then p is generically stable as a measure if and only if it is generically stable as a type. More generally, an average $\sum_{i < \omega} a_i p_i$ of types is generically stable if and only if all the p_i 's are generically stable (assuming of course $a_i > 0$ for all i).
- Any Borel probability measure on \mathbb{R} induces a smooth Keisler measure. See [42]. In fact, more generally, under mild assumptions, if M is a model equipped with a σ -algebra making externally definable sets measurable and μ_0 is a σ -additive measure on M , then μ_0 induces a generically stable Keisler measure. See [87, Theorem 5.1] for a precise statement and a proof.
- The average measure of an indiscernible sequence $I = (a_i : i \in [0, 1])$ is always generically stable. It is easy for example to check fim. It is smooth if and only if the sequence I satisfies a property called *distality*. See [86].
- Let $M = \prod_{\mathcal{D}} M_n$ be an ultraproduct of finite structures. The measure μ on M defined as the ultraproduct of the counting measures is generically stable. This follows from the VC-theorem which implies fim for the limit measure μ .

Historical Remarks

Keisler measures were introduced in the context of NIP theories by Keisler in [55]. There, he defines a notion of *smooth measure* which is weaker than the one we consider here. It says in some sense that the measure has a unique extension to the *unstable part* of the structure. In [46], Hrushovski, Peterzil and Pillay revive those ideas and use measures to study groups definable in o-minimal theories; in particular, the concept of an *invariant measure* of a group

is central in their approach. The boundedness properties come from that work. Those ideas are developed further in [47], where in particular Borel definability is observed. Finally, generically stable measures are defined and studied in [42].

Itai Ben Yaacov proves in [97] that the randomization of an NIP theory is NIP. That statement is equivalent to the following: Let $\omega(x_1, \dots)$ be a measure such that for any formula $\theta(x_1, \dots, x_n)$, and tuples $i_1 < \dots < i_n$, $j_1 < \dots < j_n$, we have $\omega(\theta(x_{i_1}, \dots, x_{i_n})) = \omega(\theta(x_{j_1}, \dots, x_{j_n}))$. Then for any formula $\phi(x; b)$ and ϵ , there can be only finitely many indices n for which $|\omega(\phi(x_n; b)) - \omega(\phi(x_{n+1}; b))| \geq \epsilon$. This result is thus the analog of Lemma 2.5 for measures and could be used to translate in a uniform way proofs from types to measures. We have not used it in order to keep this text self-contained.

Appendix A

Examples of NIP structures

A.1 Linear orders and trees

Linear orders

By a *colored order* we mean a structure $(M; <, C_i : i < \alpha)$ where the C_i 's are unary predicates and where $<$ defines a linear order on M .

Lemma A.1 (Rubin). *Let $(M; <, C_i : i < \alpha)$ be a colored order and let $a \in M$. Let $\bar{b}_1 \in M$ be a tuple of points all smaller than a and $\bar{b}_2 \in M$ a tuple of points greater than a , then $\text{tp}(\bar{b}_1/a) \cup \text{tp}(\bar{b}_2/a) \vdash \text{tp}(\bar{b}_1, \bar{b}_2/a)$.*

Proof. The proof is a straightforward back-and-forth.

Assume M is ω -saturated and we are given two pairs of tuples (\bar{b}_1, \bar{b}_2) and (\bar{b}'_1, \bar{b}'_2) satisfying the hypothesis and such that $\text{tp}(\bar{b}_1/a) = \text{tp}(\bar{b}'_1/a)$ and $\text{tp}(\bar{b}_2/a) = \text{tp}(\bar{b}'_2/a)$. Notice that necessarily $\text{tp}_{qf}(\bar{b}_1, \bar{b}_2, a) = \text{tp}_{qf}(\bar{b}'_1, \bar{b}'_2, a)$ (quantifier free types). Assume we are given $b \in M$, without loss, $b < a$. We can find $b' \in M$ such that $\text{tp}(\bar{b}_1 \hat{\ } b/a) = \text{tp}(\bar{b}'_1 \hat{\ } b'/a)$. Thus the back-and-forth goes through. \square

Proposition A.2. *Any colored order $(M; <, C_i : i < \alpha)$ is dp-minimal (in particular NIP).*

Proof. Let $a \in M$ be a singleton and $I = (\bar{a}_i : i \in \mathbb{Q})$ be an indiscernible sequence of n -tuples. For each index i write $\bar{a}_i = (a_i^1, \dots, a_i^n)$. We say that I is *cut by a* if there is $k \leq n$ and $i, i' \in \mathbb{Q}$ such that $a_i^k \leq a \leq a_{i'}^k$.

Assume that I is not cut by a , then it follows from Lemma A.1 that for any $i < i' \in \mathbb{Q}$, $\text{tp}(a/\bar{a}_i \hat{\ } \bar{a}_{i'}) \cup \text{tp}((a_j : i < j < i')/\bar{a}_i \hat{\ } \bar{a}_{i'}) \vdash \text{tp}(a + (a_j : i <$

$j < i')/\bar{a}_i \hat{=} \bar{a}_{i'}$). As $(a_j : i < j < i')$ is indiscernible over $\bar{a}_i \hat{=} \bar{a}_{i'}$, it follows that $(a_j : i < j < i')$ is indiscernible over a , and therefore I is indiscernible over a .

If now I and J are two mutually indiscernible sequences, then a singleton a cannot cut both I and J (otherwise either some point of I would cut J or some point in J would cut I). Therefore at least one of I and J is indiscernible over a . It follows that the theory is dp-minimal. \square

Trees

Definition A.3 (Tree). A tree is a structure (M, \leq) such that \leq defines a partial order on M and for every $a \in M$, the set $\{x \in M : x \leq a\}$ is linearly ordered by \leq .

A tree M is connected if for every $a, b \in M$ the set $\{x \in M : x \leq a \wedge x \leq b\}$ is non-empty.

We say that (M, \leq) is a meet-tree if it is a connected tree and in addition, for every two points $a, b \in M$, the set $\{x \in M : x \leq a \wedge x \leq b\}$ has a greatest element, which we write $a \wedge b$.

Let (M, \leq) be a meet-tree and $c \in M$ is a point. The *closed cone* of center c is by definition the set $C(c) := \{x \in M : x \geq c\}$. We can define on $C(c)$ a relation E_c by: $xE_c y$ if $x \wedge y > c$. One can easily check that this is an equivalence relation. We define an *open cone* of center c to be an equivalence class under the relation E_c .

A *leaf* of the tree M , is point in M which is maximal.

Let T_{dt} be the theory of dense meet-trees in the language $\{\leq, \wedge\}$ defined by saying that:

- \leq defines a meet-tree on the universe, and \wedge is defined as above;
- for any point x , $\{y : y \leq x\}$ is dense with no first element;
- for any point x , there are infinitely many open cones of center x .

This is a complete theory which admits elimination of quantifiers.

Lemma A.4. *Let (M, \leq) be a tree, $a \in M$, and let C denote the closed cone of center a . Let $\bar{x} = (x^1, \dots, x^n) \in (M \setminus C)^n$ and $\bar{y} = (y^1, \dots, y^m) \in C^m$. Then $tp(\bar{x}/a) \cup tp(\bar{y}/a) \vdash tp(\bar{x} \cup \bar{y}/a)$.*

Proof. The proof is back-and-forth as for Lemma A.1. \square

Proposition A.5. *Any tree (M, \leq) is dp-minimal (hence NIP).*

The proof is similar to that of Proposition A.2, but slightly longer as we have to consider various cases. We omit it here and refer the reader to [88, Proposition 4.7].

O-minimal and C-minimal structures

Let $(M; <, \dots)$ be a structure on which $<$ defines a linear order. We say that M is *o-minimal* if every definable set of M (in dimension 1) is a finite union of intervals (closed or open). One can prove that this implies that any model of $Th(M)$ is o-minimal.

As mentioned in the introduction, o-minimality is a major area of research within model theory, as it is both a powerful and versatile framework. This condition was isolated by van den Dries [90] and first systematically studied by Pillay and Steinhorn in [67]. The authors showed the fundamental *cell decomposition* theorem which states that definable subsets in any dimension can be decomposed as a finite union of *cells* (for examples, cells in dimension 1 are intervals). Thus o-minimal theories are geometrically tame: pathologies such as space-filling curves do not exist in them. An important landmark in the subject is the proof by Wilkie that \mathbb{R}_{\exp} , the reals with a predicate for the exponential function, is o-minimal. Van den Dries [91] (and later Denef and van den Dries [27]) proved that expanding the reals by adding analytic functions restricted to compact sets also yields an o-minimal structure.

Theorem A.6. *Any o-minimal theory is dp-minimal.*

Proof. Let $(M; <, \dots)$ be an o-minimal structure. Let $I = (\bar{b}_i : i \in \mathbb{Q})$ be an indiscernible sequence of n -tuples and let $\phi(x; \bar{y})$ be a formula, where x is a single variable. Then by o-minimality, for every $i \in \mathbb{Q}$, $\phi(x; \bar{b}_i)$ is a finite union of intervals. Note that the end points of those intervals are definable over b_i . By indiscernability of the sequence $(b_i : i \in \mathbb{Q})$, the number of those intervals and their type (open, closed) is constant as i varies; so are the functions sending b_i to each of the end-points of those intervals. In other words, there is a number N , \emptyset -definable functions $(g_k(\bar{y}) : k < N)$ and a quantifier free formula $\theta(x; \bar{z})$ in the language $\{=, <\}$ such that for each $i \in \mathbb{Q}$, the formula $\phi(x; \bar{b}_i)$ is equivalent to $\theta(x; g_0(\bar{y}), \dots, g_{N-1}(\bar{y}))$.

Let $a \in M$ be a singleton. If the sequence $(\text{tp}(\bar{b}_i/a) : i \in \mathbb{Q})$ is not constant, it follows from the previous analysis that a cuts the indiscernible sequence $(\text{dcl}(\bar{b}_i) : i \in \mathbb{Q})$ (where *cut* is defined as in the proof of A.2). If $I = (\bar{b}_i : i \in \mathbb{Q})$ and $J = (\bar{d}_i : i \in \mathbb{Q})$ are mutually indiscernible, then the

sequences $(\text{dcl}(\bar{b}_i) : i \in \mathbb{Q})$ and $(\text{dcl}(\bar{d}_i) : i \in \mathbb{Q})$ are mutually indiscernible. Therefore a singleton a cannot cut both. This shows that the theory is dp-minimal. \square

For more information about o-minimal theories, we refer the reader to van den Dries' book [93] and Wilkie's survey [96].

The structure $(M; <, \dots)$ is *weakly o-minimal* if in every elementary extension of M , every definable unary set is a finite union of convex sets. Adapting the proof of Theorem A.6, one can show that any weakly o-minimal structure is dp-minimal.

We now present C -minimality, which is the analogue of o-minimality for trees instead of linear orders.

Let $(M; \leq)$ be a meet-tree, and let B_M be the set of maximal branches of M , i.e., the set of maximal totally ordered subsets of M . Note that if $m \in M$ is a leaf of M then the set $\{x \in M : x \leq m\}$ is an element of B_M . Hence the set of leaves of M naturally embeds into B_M . Given $a \in B_M$ and $m \in M$, we write $m < a$ if m is in the branch a . Then $(M \cup \{B_M\}, \leq)$ is a meet-tree, the leaves of which are exactly the elements of B_M .

If $a, b, c \in B_M$, we write $C(a, b, c)$ if either $b = c \neq a$ or a, b, c are distinct and $a \wedge c < b \wedge c$. Then the structure (B_M, C) satisfies the following axioms:

- (C1) $\forall x, y, z [C(x, y, z) \rightarrow C(x, z, y)]$;
- (C2) $\forall x, y, z [C(x, y, z) \rightarrow \neg C(y, x, z)]$;
- (C3) $\forall x, y, z, w [C(x, y, z) \rightarrow (C(w, y, z) \vee C(x, w, z))]$;
- (C4) $\forall x, y [x \neq y \rightarrow \exists z \neq y C(x, y, z)]$.

Conversely, one can show that given a structure $(M; C)$ where $C(x, y, z)$ satisfies axioms (C1) - (C4), then there is a meet tree $(T; \leq)$ for which M is the set of branches of T and the C -structure on M coincides with the one defined above. Such a structure is called a C -structure.

A structure $(M; C, \dots)$ is C -minimal if any definable unary subset is equivalent to a boolean combination of formulas of the form $C(x, a, b)$, $a, b \in M$. A theory is C -minimal if all of its models are C -minimal.

Theorem A.7. *Any C -minimal theory is dp-minimal, hence NIP.*

The proof is essentially the same as that of Theorem A.6 using the fact that the formula $C(x; yz)$ has VC-dimension 1.

C -minimality was defined by Macpherson and Steinhorn in [59]; a cell decomposition result was shown by Haskell and Macpherson in [39]. Further

work has been done by Delon, Maalouf, Simonetta. The main example of a C -minimal structure is that of algebraically closed valued fields (see below).

Both o-minimality and C -minimality are subsumed in the notion of VC-minimal theory as defined by Adler in [3]. See also [34] and [25].

A.2 Valued fields

Ordered abelian groups

An ordered abelian group is a structure $(\Gamma; 0, +, <)$ such that $(\Gamma; 0, +)$ is an abelian group on which $<$ defines a linear order with the following compatibility condition:

$$x < y \implies x + z < y + z, \text{ for all } x, y, z.$$

A quantifier elimination result for ordered abelian groups is proved by Schmitt in [75]. A similar result is proved by Cluckers and Halupczok in [24]. Both languages are somewhat complicated, so we do not give the details here. We however mention two important special cases.

1. Let T_{doag} be the theory of (non-trivial) *divisible* ordered abelian group (*i.e.*, we add to the theory of ordered abelian groups the axiom $(\exists x)x \neq 0$ and for every integer n , the axiom $(\forall x)(\exists y)n \cdot y = x$. The ordering ensures that this y is unique.) Then T_{doag} is a complete theory and admits elimination of quantifiers in the language $\{0, +, <\}$.

2. Let T_{pres} be Presburger arithmetic, namely the theory of $(\mathbb{Z}; 0, 1, +, <)$. That theory does not admit elimination of quantifiers, but it does if we add predicates $\{P_n : n < \omega\}$ defined by $P_n(x) \leftrightarrow (\exists y)n \cdot y = x$.

Using the general quantifier elimination result, Gurevich and Schmitt prove in [36] the following theorem.

Theorem A.8. *Any ordered abelian group $(\Gamma; 0, +, <)$ is NIP.*

Valued fields

Since the seminal work of Ax-Kochen and Eršov, valued fields have played a important role in algebraic model theory. They also provide interesting examples of NIP structures.

We will recall briefly some facts about valuations, but we assume some familiarity with this notion in the proofs. The reader is referred to [29] for more details.

Let Γ be an ordered abelian group. We let ∞ be an additional formal element and extend the ordering and composition law on $\Gamma \cup \{\infty\}$ by declaring that ∞ is greater than all elements of Γ and setting $\gamma + \infty = \infty + \gamma = \infty + \infty = \infty$ for all $\gamma \in \Gamma$.

A valued field with value group Γ is a field K equipped with a surjective map $v : K \rightarrow \Gamma \cup \{\infty\}$ satisfying:

$$\begin{aligned} v(x) = \infty &\iff x = 0; \\ v(x + y) &\geq \min(v(x), v(y)); \\ v(xy) &= v(x) + v(y). \end{aligned}$$

We write K_v to denote the field K equipped with the valuation v .

The *valuation ring* \mathcal{O} of K_v is the ring $\{x \in K : v(x) \geq 0\}$. It is a local ring, *i.e.*, has a unique maximal ideal $\mathfrak{M} = \{x \in K : v(x) > 0\}$. The quotient $\mathbf{k} = \mathcal{O}/\mathfrak{M}$ is called the *residue field* of K_v . We let $\text{res} : \mathcal{O} \rightarrow \mathbf{k}$ be the canonical projection (called the *residue map*).

A valued field can be considered in various languages. The language L_{div} is a one sorted language $(K; 0, 1, +, -, \cdot, |)$ containing the ring language on K and a binary predicate $|$ interpreted as $x|y \iff v(x) \leq v(y)$.

One can also consider valued fields in the three sorted language L_{res} having as sorts K and \mathbf{k} equipped with their respective ring structures and Γ equipped with its ordered abelian group structure. We also have two function symbols between sorts: $v : K^* \rightarrow \Gamma$ and $\text{res} : \mathcal{O} \rightarrow \mathbf{k}$ interpreted as the valuation and residue map. (If the reader is bothered by the fact that those functions are not defined everywhere, she can extend them formally by setting $v(0) = 0$ and $\text{res}(x) = 0$ when $v(x) < 0$.)

One can define on any valued field a natural C -structure by

$$C(x, y, z) \equiv v(x - z) < v(y - z).$$

ACVF

Let ACVF denote the theory of algebraically closed (non-trivially) valued fields in the language L_{div} . The fact that the valued field K_v is algebraically closed forces the residue field \mathbf{k} to be also algebraically closed and the value group Γ to be a divisible group.

We let $\text{ACVF}_{(0,0)}$, $\text{ACVF}_{(0,p)}$ and $\text{ACVF}_{(p,p)}$ denote the theories of algebraically closed valued fields where the pair (characteristic of K_v , characteristic of \mathbf{k}) is respectively equal to $(0,0)$, $(0,p)$ and (p,p) .

Theorem A.9. *The theory ACVF eliminates quantifiers in the language L_{div} . Its completion are the theories $\text{ACVF}_{(0,0)}$, $\text{ACVF}_{(0,p)}$ and $\text{ACVF}_{(p,p)}$, for p a prime number.*

This was essentially proved by Robinson in [73]. See also [19]. The following easily follows from the theorem.

Theorem A.10. *Equipped with its natural C -structure, the theory ACVF is C -minimal, in particular it is dp -minimal, hence NIP.*

Henselian valued fields

An *angular component* is a map $\text{ac} : K_v \rightarrow \mathbf{k}$ which satisfies:

- $\text{ac}(0) = 0$;
- the restriction of ac to K_v^* has image in \mathbf{k}^* and is a morphism of multiplicative groups;
- for any $x \in K_v$ of valuation 0, $\text{ac}(x) = \text{res}(x)$.

If $x, y \in K$, then $v(x) < v(y)$ implies $v(x+y) = v(x)$ and $\text{ac}(x+y) = \text{ac}(x)$. If $v(x) = v(y)$, then either $\text{ac}(x) \neq -\text{ac}(y)$, in which case $v(x+y) = v(x)$ and $\text{ac}(x+y) = \text{ac}(x) + \text{ac}(y)$, or $\text{ac}(x) = -\text{ac}(y)$ in which case $v(x+y) > v(x)$ and we cannot say anything about $\text{ac}(x+y)$.

Fact A.11. *If K_v is an ω_1 -saturated valued field, then K_v admits an angular component.*

See for example [20]. Keeping this fact in mind, we may restrict our study to valued fields with an angular component. Let L_{Pas} be the language $L_{\text{res}} \cup \{\text{ac}\}$ where ac is a new function symbol from K to \mathbf{k} .

Theorem A.12 (Pas [65]). *The (incomplete) theory of henselian valued field of residue characteristic 0 eliminates fields quantifiers in the language L_{Pas} .*

See also [19] for a proof. Examples of such fields include $\mathbb{C}((t))$, $\mathbb{R}((t))$: the fields of Laurent series respectively over \mathbb{R} , or \mathbb{C} , or more generally any field of the type $k((\Gamma))$ where k is a field of characteristic 0 and Γ an ordered abelian group (see [29, Exercise 3.5.6]).

A direct consequence is the celebrated result of Ax-Kochen and Eršov (independently) on elementary equivalence of henselian valued fields.

Theorem A.13 (Ax-Kochen, Eršov). *Let (K, v) and (L, w) be two henselian valued fields. Denote by Γ_K, Γ_L their respective value groups equipped with the ordered group structure and by \mathbf{k}, \mathbf{l} the residue fields in the ring language. Assume that \mathbf{k} and \mathbf{l} have characteristic zero. Then we have:*

$$K_v \equiv L_w \iff (\Gamma_K \equiv \Gamma_L \text{ and } \mathbf{k} \equiv \mathbf{l}).$$

In the same spirit, if K_v is independent, then this can be traced down to the residue field (recall that the value group, being a pure ordered abelian group, is always NIP).

Theorem A.14 (Delon [26]). *A Henselian field of residue characteristic 0 is NIP if and only if both its value group and its residue field are NIP.*

Knowing Theorem A.8, we do not need to mention the value group.

Corollary A.15. *A Henselian field of residue characteristic 0 is NIP if and only if its residue field is NIP.*

Delon's proof uses a coheir-counting argument and requires to first understand types. We sketch here a more direct (though not necessarily simpler) argument using indiscernible sequences. As discussed previously, we may assume that our field K_v is equipped with an angular component ac , and we work in the language L_{Pas} (the reduct of an NIP theory is NIP, so the result will follow for the languages L_{div} and L_{res} .)

Lemma A.16. *Let $P(X) = \sum_{k \leq n} a_k X^k$ be in $K_v[X]$ and let $(x_i : i < \omega)$ be a sequence of elements of K_v such that $(v(x_i))_{i < \omega}$ is monotonous (increasing or decreasing), then there are $r \leq n$ and $t < \omega$ such that for all $i \geq t$ and $k \neq r$, $v(P(x_i)) = v(a_r x_i^r) < v(a_k x_i^k)$ (hence also $\text{ac}(P(x_i)) = \text{ac}(a_r x_i^r)$).*

Proof. As the sequence $(v(x_i))_{i < \omega}$ is increasing, there is $t < \omega$ such that for $i \geq t$, $v(x_i)$ has the same relative position with respect to all values of the form $\frac{v(a_k) - v(a_{k'})}{k' - k}$, for $k, k' \in \Gamma$ (if Γ is not divisible, those elements live in its divisible hull, to which the order extends in a unique way). In particular, $v(x_i)$ is not equal to any of those values.

Hence for $i \geq t$, the elements $v(a_k) + k \cdot v(x_i)$ for $k \leq n$ are pairwise distinct, and there is $r \leq n$ not depending on i such that their minimum is $v(a_r) + r \cdot v(x_i)$. \square

We now classify indiscernible sequences of singletons depending on their type in the C -structure. Let $(x_i : i < \omega)$ be an indiscernible sequence of singletons of K_v .

Case 0: the sequence $(v(x_i) : i < \omega)$ is non-constant. Then by indiscernability, it is either decreasing or increasing.

We now assume that the sequence $(v(x_i) : i < \omega)$ is constant. For $0 < i < \omega$, let $y_i = x_i - x_0$. Note that it is not possible for the sequence $(v(y_i) : 0 < i < \omega)$ to be increasing. For then, we would have for $i > 1$, $v(x_i - x_1) = v((y_i) - v(y_1)) = v(y_1)$. Thus the sequence $v(x_i - x_1)$ would be constant, while the sequence $v(x_i - x_0)$ is not and this contradicts indiscernibility.

Case I: The sequence $(v(y_i) : 0 < i < \omega)$ is decreasing.

Case II: The sequence $(v(y_i) : 0 < i < \omega)$ is constant and the sequence $(\text{ac}(y_i) : 0 < i < \omega)$ is not constant.

Case III: The sequences $(v(y_i) : 0 < i < \omega)$ and $(\text{ac}(y_i) : 0 < i < \omega)$ are both constant. Then we have $v(x_2 - x_1) = v(y_2 - y_1) > v(x_2 - x_0)$. Let x_ω be such that $(x_i : i \leq \omega)$ is indiscernible. For $i < \omega$, we let $z_i = x_i - x_\omega$. Then the sequence $(v(z_i) : i < \omega)$ is decreasing.

Lemma A.17. *Let $(x_i : i < \omega)$ be an indiscernible sequence of singletons of K_v . Then there are:*

- an indiscernible sequence $(\alpha_i : i < \omega)$ of elements of Γ ,
 - an indiscernible sequence $(b_i : i < \omega)$ of elements of \mathbf{k} , such that:
for any $P(X) \in K_v[X]$, there is $r < \omega$, and $\gamma \in \Gamma$ such that $v(P(x_i)) = \gamma + r \cdot \alpha_i$ for all i large enough.
- Also, there is $Q \in \mathbf{k}[X]$ such that for all i large enough, $\text{ac}(P(x_i)) = Q(b_i)$.

Proof. If the sequence $(x_i)_{i < \omega}$ falls in case 0, we are done by Lemma A.16. Assume it falls in case I. There is a polynomial $DP(Y) \in K_v[Y]$ such that we have $P(x_0 + Y) = P(x_0) + DP(Y)$. Then for all $0 < i < \omega$, we have $P(x_i) = P(x_0) + DP(y_i)$. We conclude by applying Lemma A.16 to the sequence $(y_i)_{0 < i < \omega}$ and the polynomial $P(x_0) + DP(Y)$.

If $(x_i)_{i < \omega}$ falls in case III, we write similarly $P(x_\omega + Z) = P(x_\omega) + DP_1(Z)$ and apply Lemma A.16 to the sequence $(z_i)_{i < \omega}$ and the polynomial $DP_1(Z)$.

Finally, assume that we are in case II. Write $P(X) = \sum_{k < n} a_k X^k$. Let $v_0 = v(x_0)$. Let $A \subset n$ be the set of $k < n$ for which $v(a_k) + k \cdot v_0$ is minimal. Let $q(t) \in \mathbf{k}[t]$ be the polynomial $\sum_{k \in A} \text{ac}(a_k) t^k$. Then for some $i_* < \omega$, for every $i_* < i < \omega$, $\text{ac}(a_i)$ is not a root of $q(t)$. For such an i , we have $v(P(x_i)) = v_0$ and $\text{ac}(P(x_i)) = q(\text{ac}(x_i))$. The lemma follows. \square

We now prove Theorem A.14.

By Theorem A.12 (and keeping Proposition 2.14 in mind), it is enough to show that the following formulas are NIP:

- $\phi(x, \bar{y}) = 0$, where x, \bar{y} are variables of sort K_v and ϕ is a quantifier free formula in the ring language of K_v ;
- $\psi(x, \bar{t}(\bar{y}))$ where x is a variable of sort Γ , \bar{y} variables of sorts K_v and Γ , ψ is a formula in the language of Γ , and \bar{t} is a tuple of terms with image in Γ ;
- $\theta(x, \bar{t}(\bar{y}))$ where x is a variables of the sort \mathbf{k} , \bar{y} variables from K_v and \mathbf{k} , θ is a formula in the language of \mathbf{k} , and \bar{t} is a tuple of terms with image in \mathbf{k} ;
- $\psi(v(P_1(x, \bar{y}_1)), \dots, v(P_n(x, \bar{y}_1)), \bar{y}_2)$ where \bar{y}_2 are variables from Γ and ψ is a formula in the language of Γ ;
- $\theta(\text{ac}(P_1(x, \bar{y}_1)), \dots, \text{ac}(P_n(x, \bar{y}_1)), \bar{y}_2)$, where \bar{y}_2 are variables from \mathbf{k} and θ is a formula in the language of \mathbf{k} .

In the first three cases, the results follows from the fact that algebraically closed fields, $Th(\Gamma)$ and $Th(\mathbf{k})$ respectively are NIP.

Assume that the formula $\varphi(x; \bar{y}_1, \bar{y}_2) = \psi(v(P_1(x, \bar{y}_1)), \dots, v(P_n(x, \bar{y}_1)), \bar{y}_2)$ has IP. Then there is an indiscernible sequence $(x_i : i < \omega)$ of singletons and parameters \bar{b}_1, \bar{b}_2 such that $\varphi(x_i; \bar{b}_1, \bar{b}_2)$ holds if and only if i is even.

By Lemma A.17, there is an indiscernible sequence $(\alpha_i : i < \omega)$ of elements of Γ and for $k \leq n$, $r_n < \omega$ and $\gamma_n \in \Gamma$ such that $v(P_k(x_i, \bar{b}_1)) = \gamma_k + r_k \cdot \alpha_i$ for all i large enough. Hence we can replace each $v(P_1(x, \bar{y}_1))$ in the formula ψ by a term in the language of Γ and we obtain a contradiction to the fact that the ordered abelian group Γ is NIP.

We treat similarly the last case. This finishes the proof of Theorem A.14.

The theory ACVF has been intensely studied in the last ten years. Haskell, Hrushovski and Macpherson provided in [37] a description of imaginaries. They showed in [38] how types can be decomposed into an o-minimal component coming from the value group and a stable *quotient*, internal to the residue field. This property is referred to as *metastability*. A measure-theoretic analog of the Ax-Kochen principle is studied in [44] by Hrushovski and Kazhdan. Most recently, Hrushovski and Loeser [45] have given a model-theoretic construction of Berkovich spaces from rigid geometry.

The p -adics

Let \mathbb{Q}_p denote the usual field of p -adic numbers. The valued field \mathbb{Q}_p is henselian, but of residue characteristic p , hence the previous results do not

apply. Nonetheless, its theory is well understood as we explain now.

First we give an axiomatization of $Th(\mathbb{Q}_p)$ in the language L_{res} . In addition to the paper [10] by Ax and Kochen where the result first appeared, we refer the reader to Cherlin's [20, II Th. 40].

Theorem A.18. *Let T_p be the following theory, in the language L_{res} :*

- *Axioms for henselian valued fields of characteristic zero.*
- *The residue field \mathbf{k} is isomorphic to \mathbb{F}_p .*
- *The value group Γ satisfies the theory of $(\mathbb{Z}; 0, +, <)$.*
- *$v(p) = 1$ is the smallest positive element of Γ .*

Then T_p is a complete theory, and is equal to the theory of \mathbb{Q}_p in the language L_{res} .

It is worth noting that in the field \mathbb{Q}_p , the valuation is definable in the pure field structure (namely, the valuation ring is definable by the formula $\phi(x) \equiv \exists t, t^2 = 1 + p^3 x^4$), hence we also obtain an axiomatization of $Th(\mathbb{Q}_p)$ as a pure field.

The valued field \mathbb{Q}_p does not admit elimination of field quantifiers in either L_{div} or L_{res} . To remedy that, we present two enriched languages. First, let L_{mac} be the language L_{div} to which we add predicates $\{P_n : 0 < n < \omega\}$. The valued field \mathbb{Q}_p is made into an L_{mac} -structure by interpreting P_n as the set of n -th powers.

Theorem A.19 (Macintyre [58]). *The field \mathbb{Q}_p has elimination of quantifiers in the language L_{mac} .*

Another way to obtain quantifier elimination is by adding angular components as we did in the case of residue characteristic zero. However, we now need to add infinitely many. Let K_v be a valued field of characteristic 0, whose residue field has characteristic p . A family $(ac_n : n < \omega)$ is a compatible family of angular components if:

- ac_n is a map $K_v \rightarrow \mathcal{O}/p^n \mathcal{O}$;
- $ac_n(0) = 0$; the restriction of ac_n to K_v^* has image in $(\mathcal{O}/p^n \mathcal{O})^*$ and is a morphism of groups;
- let π_n denote the canonical projection $\pi_n : \mathcal{O}/p^{n+1} \mathcal{O} \rightarrow \mathcal{O}/p^n \mathcal{O}$, then $\pi_n \circ ac_{n+1} = ac_n$.

Let $L_{Pas, \omega}$ be the language $L_{res} \cup \{ac_n : n < \omega\}$.

We can explicitly construct a family of compatible angular components on \mathbb{Q}_p . Let $x \in \mathbb{Q}_p^*$ and let $\mathbf{v} = v(x)$. Set $x_0 = x/p^{\mathbf{v}}$. Then $x_0 \in \mathcal{O} \setminus p\mathcal{O}$. Now let $ac_n(x)$ be the image of x_0 in $\mathcal{O}/p^n \mathcal{O}$.

The properties of the sequence $(ac_n : 0 < n < \omega)$ are easy to check.

Theorem A.20 (Belair [14]). *The canonical expansion of \mathbb{Q}_p in the language $L_{Pas,\omega}$ admits elimination of quantifiers.*

Now one can adapt the proof of the previous subsection to show that \mathbb{Q}_p is NIP (or alternatively adapt Delon's proof, which is what Belair does in [14]). Actually, more is true.

Theorem A.21 (Dolich, Goodrick, Lippel [28]). *For any prime p , the field \mathbb{Q}_p of p -adics in the language L_{div} is dp -minimal.*

Appendix B

Stability

Stability is the first, and most important, dividing line discovered by Shelah in his work on classification theory. Stable theories have a canonical notion of independence (non-forking) from which a dimension theory and strong structure results follow. It has proved to be a fundamental and extremely rich property, despite the fact that stable structures are rare. In algebra, the main examples are algebraically closed fields, separably closed fields and differentially closed fields.

We only record here basic facts about stability which are used in the text. The reader can consult [66] or [72] for more information.

Definition B.1 (Stable formula). A formula $\phi(x; y)$ is stable if there does not exist points $(a_i b_i : i < \omega)$ such that:

$$\models \phi(a_i; b_j) \iff i \leq j.$$

If a formula $\phi(x; y)$ is unstable, then by Ramsey, we can assume that the sequence $(a_i b_i : i < \omega)$ witnessing instability is indiscernible. We easily obtain the following equivalent characterization.

Lemma B.2. *A formula $\phi(x; y)$ is stable if and only if there is no indiscernible sequence $(a_i : i < \omega^2)$ and $b \in \mathcal{U}$ such that:*

$$\phi(a_i; b) \iff i < \omega.$$

There are a number of equivalent characterizations of stability.

If $A \subset \mathcal{U}$ and $p \in S_\phi(A)$ is a ϕ -type over A , we say that p is *definable* if there is a formula $d\phi(y; z)$ and a tuple $b \in A$ such that for all $a \in A$ we have $p \vdash \phi(x; a) \iff \models d\phi(a; b)$.

Proposition B.3. *Let $\phi(x; y)$ be a formula. Then the following are equivalent:*

- (i) $\phi(x; y)$ is stable;
- (ii) for all A infinite, we have $|S_\phi(A)| \leq |A|$;
- (iii) there is an integer k such that for all A , finite or infinite, we have $|S_\phi(A)| \leq |A|^k$;
- (iv) for all A , all ϕ -types over A are definable.

Note that easily (iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i). The implication (i) \Rightarrow (iv) is more complicated, see [66, Lemma 2.2]. Moreover, if $A = M$ is a model, then every ϕ -type over M can be defined by a ϕ -formula, *i.e.*, $d\phi$ can be taken to be a boolean combination of formulas of the form $\phi(x; b)$ for $b \in M$.

Appendix C

Probability theory

We recall here some basic results of probability theory that we need in this text. We refer the reader to any introductory book on the subject for more details. Alon and Spencer's book [6], although not a textbook on probability theory, is a nice reference.

A *probability space* (X, \mathcal{B}, μ) is a set X equipped with a σ -algebra \mathcal{B} and a σ -additive measure μ on \mathcal{B} such that $\mu(A) = 1$. For each integer k , the cartesian power X^k is naturally equipped with the product σ -algebra $\mathcal{B}^{\otimes k}$ which is the σ -algebra generated by sets of the form $B_1 \times \cdots \times B_k$ for $B_1, \dots, B_k \in \mathcal{B}$. The product measure μ^k is defined as the unique probability measure on $(X^k, \mathcal{B}^{\otimes k})$ such that $\mu^k(B_1 \times \cdots \times B_k) = \mu(B_1) \cdots \mu(B_k)$.

If $A \subset X$ is a measurable subset of X , we let $\mathbf{1}_A$ be its characteristic function. We write $\text{Prob}(A) = \mu(A)$. If $f, g : X \rightarrow \mathbb{R}$ are measurable functions, we will write $\{f \geq g\}$ for $\{x \in X : f(x) \geq g(x)\}$ and $\text{Prob}(f \geq g)$ instead of $\text{Prob}(\{x \in X : f(x) \geq g(x)\})$.

A measurable function $f : X \rightarrow \mathbb{R}$ is called a *random variable*. Its expectancy $E(f)$ is defined as $\int f(x) d\mu$. Its variance is defined as $\text{Var}(f) = E((f - E(f))^2)$.

Two random variables $f, g : X \rightarrow \mathbb{R}$ are *uncorrelated* if $E(fg) = E(f)E(g)$. If f, g are uncorrelated random variables and $a, b \in \mathbb{R}$, then we have $\text{Var}(a \cdot f + b \cdot g) = a^2 \text{Var}(f) + b^2 \text{Var}(g)$.

Random variables $f_1, \dots, f_n : X \rightarrow \mathbb{R}$ are said to be *independent* if for any measurable sets B_1, \dots, B_n of \mathbb{R} , we have $\text{Prob}(\bigcap \{f_k \in B_k\}) = \prod \text{Prob}(f_k \in B_k)$. In particular, two independent variables are uncorrelated.

Lemma C.1 (First moment). *Let $f : X \rightarrow \mathbb{R}^+$ be a random variable. Then for all $r > 0$, we have*

$$\text{Prob}(f \geq r) \leq \frac{\mathbb{E}(f)}{r}.$$

Proof. Write that $\mathbb{E}(f) \leq \mathbb{E}(f \cdot \mathbf{1}_{\{f \geq r\}}) \leq r \text{Prob}(f \geq r)$. \square

Proposition C.2 (Chebyshev's inequality). *Let $f : X \rightarrow \mathbb{R}$ be measurable. Then we have*

$$\text{Prob}(|f - \mathbb{E}(f)| \geq \epsilon) \leq \frac{\text{Var}(f)}{\epsilon^2}.$$

Proof. Apply the first moment inequality to the random variable $(f - \mathbb{E}(f))^2$ and $r = \epsilon^2$. \square

Proposition C.3 (Weak law of large numbers). *Let $A \subseteq X$ be a measurable subset of X and fix $\epsilon > 0$, then*

$$\lim_{n \rightarrow +\infty} \mu^n \left(\left| \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{x_i \in A\}} - \mu(A) \right| \geq \epsilon \right) = 0.$$

Proof. Fix some integer n . For $i \leq n$, the random variable $\mathbf{1}_{\{x_i \in A\}} : X^n \rightarrow \mathbb{R}$ has expectancy equal to $\mu(A)$ and variance $\mu(A)(1 - \mu(A)) \leq 1$. Also for $i \neq j$, the random variables $\mathbf{1}_{\{x_i \in A\}}$ and $\mathbf{1}_{\{x_j \in A\}}$ are uncorrelated (in fact, they are independent). It follows that the random variable $\frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{x_i \in A\}}$ has expectancy $\mu(A)$ and variance $\leq 1/n$.

By Chebychev's inequality, we deduce that

$$\text{Prob} \left(\left| \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{x_i \in A\}} - \mu(A) \right| \geq \epsilon \right) \leq \frac{1}{n\epsilon^2},$$

from which the result follows. \square

Proposition C.4 (Chernoff's bound, special case). *Let f_1, \dots, f_n be independent random variables, such that $\text{Prob}(f_k = 1) = \text{Prob}(f_k = -1) = 1/2$, for all k . Let $\epsilon > 0$, then letting $g = \frac{1}{n} \sum f_k$, we have*

$$\text{Prob}(|g| \geq \epsilon) \leq 2 \exp \left(-\frac{n\epsilon^2}{2} \right).$$

For a proof, see [6, Appendix A]

Index

- ϕ -type, 10
- ϕ^0, ϕ^1 , 10
- abelian, 23
- abelian groups, 101
- ACVF, 46, 102
- $\text{alt}(\phi), \text{alt}(\phi, I)$, 15
- alternation, 15
- Baldwin-Saxl, 19
- based, 11, 24
- $\text{bdd}(A)$, 55, 60
- Borel-definable, 87, 88
- bounded equivalence relation, 54
- broom lemma, 62
- C-minimal, 100
- Chebyshev, 112
- Chernoff's bound, 112
- coheir, 21
- coheirs, 25
- compact strong-type, 55
- completion, 44
- concentrate, 10
- convex equivalence relation, 34
- critical points, 37
- $\text{ded}(\lambda)$, 17
- definable, 20, 22, 29, 30
 - measure, 87
- dividing, 56, 66
 - for measures, 86
- dp-minimal, 49, 98
- dp-rank, 42, 44
- EM-type, 10
- endless, 10
- ϵ -net, 74
- eventual type, 24
- extension base, 62
- externally definable, 32
- Farkas, 75
- finitely satisfiable, 20, 22
 - measure, 87
- first moment, 111
- forking, 53, 56, 60, 66
 - for measures, 86
- generically stable, 26
 - measure, 90, 93
- groups, 19, 23, 50, 101
- heir, 64
- henselian, 103
- honest definition, 30, 31
 - uniformity, 76
- indiscernible sequence, 10
- invariant type, 20
 - measure, 86
- IP formula, 13
- κ_{ict} , 47

- Keisler (measure), 79
- KP, 55
- large numbers (law), 112
- Lascar strong-type, 54
- Lascar strong-types, 59
- low, 67
- measures, 79
 - Borel, 81
- monster model, 9
- Morley (sequence), 22
- Morley sequence, 62
- mutually indiscernible, 39
- NIP formula, 13
- NTP_2 , 61
- o-minimal, 99
- order, 97
- p-adics, 106
- pair, 9
- (p, q) -property, 74
- (p, q) -theorem, 74
- probability, 111
- product (of types), 22
- Sauer, 71
- shatter, 70
- Shelah expansion, 33
- shrinking, 29, 34
- small, 9
- smooth, 84
- SOP (strict order property), 16
- stability, 109
- stably embedded, 32
- strictly non-forking, 64
 - sequence, 65
- strongly dependent, 48
- superstable, 48
- support, 81
- totally indiscernible, 11, 25
- tree, 98
- UDTFS, 77
- valued fields, 101
- Vapnik-Chervonenkis, 72
- VC-dimension, 69
 - for formulas, 13
- weak-stable embeddedness, 31

Bibliography

- [1] Hans Adler. Kim’s lemma in ntp2 theories. handwritten notes, available on the author’s webpage.
- [2] Hans Adler. Strong theories, burden, and weight.
- [3] Hans Adler. Theories controlled by formulas of vapnik-chervonenkis codimension 1. draft.
- [4] Hans Adler. An introduction to theories without the independence property. *Archive of Mathematical Logic*, 2008. to appear.
- [5] N. Alon and D.J. Kleitman. Piercing convex sets and the hadwiger-debrunner (p, q) problem. *Advances in Math.*, 96:103–112, 1992.
- [6] N. Alon and J.H. Spencer. *The probabilistic method*. Wiley-Interscience series in discrete mathematics and optimization. Wiley, 2000.
- [7] Noga Alon. On the density of sets of vectors. *Discrete Mathematics*, 46(2):199 – 202, 1983.
- [8] Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald Macpherson, and Sergei Starchenko. Vapnik-chervonenkis density in some theories without the independence property, i. *preprint*, 2011.
- [9] Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald Macpherson, and Sergei Starchenko. Vapnik-chervonenkis density in some theories without the independence property, ii. *preprint*, 2011.
- [10] James Ax and Simon Kochen. Diophantine problems over local fields ii. a complete set of axioms for p-adic number theory. *American Journal of Mathematics*, 87(3):pp. 631–648, 1965.

- [11] Yerzhan Baisalov and Bruno Poizat. Paires de structures o-minimales. *Journal of Symbolic Logic*, 63:570–578, 1998.
- [12] J. T. Baldwin and Jan Saxl. Logical stability in group theory. *J. Austral. Math. Soc. Ser. A*, 21(3):267–276, 1976.
- [13] John Baldwin and Michael Benedikt. Stability theory, permutations of indiscernibles, and embedded finite models. *Transactions of the American Mathematical Society*, 352(11):4937–4969, 11 2000.
- [14] Luc Bélair. Types dans les corps valués munis d’applications coefficients. *Illinois J. Math.*, 43(2):410–425, 1999.
- [15] Alexander Berenstein. Lovely pairs and dense pairs of o-minimal structures. Preprint.
- [16] Alexander Berenstein, Alfred Dolich, and Alf Onshuus. The independence property in generalized dense pairs of structures. *J. Symbolic Logic*, 76(2):391–404, 2011.
- [17] Gareth Boxall. Nip for some pair-like theories. *Archive for Mathematical Logic*, 50:353–359, 2011. 10.1007/s00153-010-0218-3.
- [18] E. Casanovas. *Simple Theories and Hyperimaginaries*. Lecture Notes in Logic. Cambridge University Press, 2011.
- [19] Zoé Chatzidakis. Théorie des modèles des corps valués. <http://www.logique.jussieu.fr/~zoe/>.
- [20] Gregory L. Cherlin. *Model-theoretic algebra: selected topics*, volume vol. 52 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1976.
- [21] Artem Chernikov and Itay Kaplan. Forking and dividing in NTP_2 theories. *Journal of Symbolic Logic*, 77(1):1–20, 2012.
- [22] Artem Chernikov and Pierre Simon. Externally definable sets and dependent pairs ii. submitted.
- [23] Artem Chernikov and Pierre Simon. Externally definable sets and dependent pairs. *Israel J. Math.*, 2012. DOI 10.1007/s11856-012-0061-9.

- [24] Raf Cluckers and Immanuel Halupczok. Quantifier elimination in ordered abelian groups. *Confluentes Mathematici*, 3:587–615, 2011.
- [25] Sarah Cotter and Sergei Starchenko. Forking in vc-minimal theories.
- [26] Françoise Delon. Types sur $\mathbf{C}((X))$. In *Study Group on Stable Theories (Bruno Poizat), Second year: 1978/79 (French)*, pages Exp. No. 5, 29. Secrétariat Math., Paris, 1981.
- [27] J. Denef and L. van den Dries. p-adic and real subanalytic sets. *The Annals of Mathematics*, 128(1):pp. 79–138, 1988.
- [28] Alfred Dolich, David Lippel, and John Goodrick. Dp-minimal theories: basic facts and examples. *Notre Dame Journal of Formal Logic*, 52(3):267–288, 2011.
- [29] A.J. Engler and A. Prestel. *Valued fields*. Springer monographs in mathematics. Springer, 2006.
- [30] Sally Floyd and Manfred Warmuth. Sample compression, learnability, and the vapnik-chervonenkis dimension. In *MACHINE LEARNING*, pages 269–304, 1995.
- [31] Peter Frankl. On the trace of finite sets. *Journal of Combinatorial Theory, Series A*, 34(1):41 – 45, 1983.
- [32] Vincent Guingona. Dependence and isolated extensions. *Proc. Am. Math. Soc.*, 139:3349–3357, 2011.
- [33] Vincent Guingona. On uniform definability of types over finite sets. *Journal of Symbolic Logic*, 77(2):499–514, 2012.
- [34] Vincent Guingona and M. C. Laskowski. On vc-minimal theories and variants. submitted.
- [35] Ayhan Günaydin and Philipp Hieronymi. Dependent pairs. *Journal of Symbolic Logic*, 76(2):377–390, 2011.
- [36] Y. Gurevich and P. H. Schmitt. The theory of ordered abelian groups does not have the independence property. *Trans. Amer. Math. Soc.*, 284(1):171–182, 1984.

- [37] Deirdre Haskell, Ehud Hrushovski, and Dugald Macpherson. Definable sets in algebraically closed valued fields. part i: elimination of imaginaries. *J. reine angew. Math.*, 597:175–236, 2006.
- [38] Deirdre Haskell, Ehud Hrushovski, and Dugald Macpherson. *Stable domination and independence in algebraically closed valued fields*, volume 30 of *Lecture Notes in Logic*. Association for Symbolic Logic, Chicago, IL, 2008.
- [39] Deirdre Haskell and Dugald Macpherson. Cell decompositions of c-minimal structures. *Annals of Pure and Applied Logic*, 66(2):113 – 162, 1994.
- [40] Wilfrid Hodges. *A Shorter Model Theory*. Cambridge University Press, 1997.
- [41] Wilfrid Hodges, A. H. Lachlan, and Saharon Shelah. Possible orderings of an indiscernible sequence. *Bulletin of the London Mathematical Society*, 9(2):212–215, 1977.
- [42] Ehud Hrushovski, Anand Pillay, and Pierre Simon. Generically stable and smooth measures in NIP theories. *Trans. Amer. Math. Soc.*, 2012. to appear.
- [43] Ehud Hrushovski, Anand Pillay, and Pierre Simon. A note on generically stable measures and *fsg* groups. *Journal of Symbolic Logic*, 2012. to appear.
- [44] Ehud Hrushovski and David Kazhdan. Integration in valued fields.
- [45] Ehud Hrushovski and Francois Loeser. Non-archimedean tame topology and stably dominated types. 2010.
- [46] Ehud Hrushovski, Ya’acov Peterzil, and Anand Pillay. Groups, measures, and the NIP. *J. Amer. Math. Soc.*, 21(2):563–596, 2008.
- [47] Ehud Hrushovski and Anand Pillay. On nip and invariant measures. *Journal of the European Mathematical Society*, 13:1005–1061, 2011.
- [48] H. R. Johnson and M. C. Laskowski. Compression schemes, stable definable families, and o-minimal structures. *Discrete Comput. Geom.*, 43(4):914–926, 2010.

- [49] Itay Kaplan, Alf Onshuus, and Alex Usvyatsov. Additivity of the dp-rank. *Trans. Amer. Math. Soc.*, 2012.
- [50] Itay Kaplan, Tom Scanlon, and Frank O. Wagner. Artin-schreier extensions in dependent and simple fields. *submitted*.
- [51] Itay Kaplan and Saharon Shelah. Chain conditions in dependent groups. Preprint.
- [52] Itay Kaplan and Saharon Shelah. A dependent theory with few indiscernibles. Preprint.
- [53] Itay Kaplan and Saharon Shelah. Examples in dependent theories. Preprint.
- [54] Itay Kaplan and Alex Usvyatsov. Strict independence in dependent theories. preprint.
- [55] H. Jerome Keisler. Measures and forking. *Ann. Pure Appl. Logic*, 34(2):119–169, 1987.
- [56] Krzysztof Krupinski. On ω -categorical groups and rings with nip. Preprint.
- [57] Michael C. Laskowski. Vapnik-chervonenkis classes of definable sets. *J. London Math. Soc.*, 45:377–384, 1992.
- [58] Angus MacIntyre. On definable subsets of p-adic fields. *The Journal of Symbolic Logic*, 41(3):pp. 605–610, 1976.
- [59] Dugald Macpherson and Charles Steinhorn. On variants of o-minimality. *Ann. Pure Appl. Logic*, 79(2):165–209, 1996.
- [60] Dugald Macpherson and Katrin Tent. Pseudofinite groups with nip theory and definability in finite simple groups. Preprint.
- [61] David Marker. *Model Theory: An Introduction*. Springer, 2002.
- [62] Jirí Matousek. Bounded vc-dimension implies a fractional helly theorem. *Discrete & Computational Geometry*, 31:251–255, 2004. 10.1007/s00454-003-2859-z.

- [63] William Mitchell. Aronszajn trees and the independence of the transfer property. *Ann. Math. Logic*, 5:21 – 46, 1972.
- [64] Alf Onshuus and Alexander Usvyatsov. On dp-minimality, strong dependence and weight. *Journal of Symbolic Logic*, 76(3):737–758, 2011.
- [65] Johan Pas. Uniform p-adic cell decomposition and local zeta functions. *J. reine angew. Math.*, 399:137 – 172, 1989.
- [66] A. Pillay. *Geometric stability theory*. Oxford logic guides. Clarendon Press, 1996.
- [67] Anand Pillay and Charles Steinhorn. Definable sets in ordered structures. i. *Transactions of the American Mathematical Society*, 295(2):pp. 565–592, 1986.
- [68] Anand Pillay and Predrag Tanović. Generic stability, regularity, and quasi-minimality. 2009.
- [69] B. Poizat. *Groupes stables*. Nur al-Mantiq wal-Ma’rifah. 1987.
- [70] Bruno Poizat. Théories instables. *J. Symbolic Logic*, 46(3):513–522, 1981.
- [71] Bruno Poizat. Post-scriptum à ”théories instables”. *The Journal of Symbolic Logic*, 48(1):pp. 60–62, 1983.
- [72] Bruno Poizat. *Cours de théorie des modèles*. Nur al-Mantiq wal-Mari’fah, 1985.
- [73] Abraham Robinson. *Complete theories*. North-Holland, Amsterdam, 1956.
- [74] N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145 – 147, 1972.
- [75] P. H. Schmitt. *Model theory of ordered abelian groups*. PhD thesis, Habilitationsschrift, 1982.
- [76] Alexander Schrijver. *Theory of Linear and Integer Programming*. Wiley, New York, 1986.
- [77] Saharon Shelah. Strongly dependent theories. Submitted.

- [78] Saharon Shelah. Stability, the f.c.p., and superstability; model theoretic properties of formulas in first order theory. *Ann. Math. Logic*, 3(3):271–362, 1971.
- [79] Saharon Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific J. Math.*, 41:247–261, 1972.
- [80] Saharon Shelah. Simple unstable theories. *Ann. Math. Logic*, 19(3):177–203, 1980.
- [81] Saharon Shelah. *Classification theory and the number of nonisomorphic models*, volume 92 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, second edition, 1990.
- [82] Saharon Shelah. Classification theory for elementary classes with the dependence property—a modest beginning. *Sci. Math. Jpn.*, 59(2):265–316, 2004. Special issue on set theory and algebraic model theory.
- [83] Saharon Shelah. Dependent first order theories, continued. *Israel Journal of Mathematics*, 173(1):1–60, 2009.
- [84] Saharon Shelah. A dependent dream and recounting types. *preprint*, 2010. 950 in Shelah archive.
- [85] Saharon Shelah. Dependent theories and the generic pair conjecture. *Communications in Contemporary Mathematics*, submitted, 2010. math.LO/0702292.
- [86] Pierre Simon. Distal and non-distal theories. 2011.
- [87] Pierre Simon. Finding generically stable measures. to appear in the Journal of Symbolic Logic, 2011.
- [88] Pierre Simon. On dp-minimal ordered structures. *Journal of Symbolic Logic*, 76(4):448–460, 2011.
- [89] Alexander Usvyatsov. On generically stable types in dependent theories. *Journal of Symbolic Logic*, 74(1):216–250, 2009.
- [90] Lou van den Dries. Remarks on tarski’s problem concerning $(\mathbf{r}, +, *, \exp)$. In G. Longo G. Lolli and A. Marcja, editors, *Logic Colloquium ’82*, volume 112 of *Studies in Logic and the Foundations of Mathematics*, pages 97 – 121. Elsevier, 1984.

- [91] Lou van den Dries. A generalization of the tarski-seideberg theorem, and some nondefinability results. *Bull. Amer. Math. Soc.*, 15:189–193, 1986.
- [92] Lou van den Dries. Dense pairs of o-minimal structures. *Fund. Math.*, 157:61–78, 1998.
- [93] L.P.D. van den Dries. *Tame Topology and O-minimal Structures*. London Mathematical Society Lecture Note Series. Cambridge University Press, 1998.
- [94] V. N. Vapnik and A. Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.*, 16(2):264–280, 1971.
- [95] Frank O. Wagner. *Simple theories*, volume 503 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2000.
- [96] Alex Wilkie. O-minimal structures. *Séminaire Bourbaki (2007/08), exp. n°985, Asérisque*, 326:131 – 142, 2009.
- [97] Itai Ben Yaacov. Continuous and random vapnik-chervonenkis classes. *Israel Journal of Mathematics*, 173:309–333, 2009.